

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner # : \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date \_\_\_\_\_

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

## STAFF USE ONLY

Searcher: Toby R.  
 Searcher Phone #: 808-3534  
 Searcher Location: \_\_\_\_\_  
 Date Searcher Picked Up: 5/23  
 Date Completed: 5/29  
 Searcher Prep & Review Time: 20  
 Clerical Prep Time: 1  
 Online Time: 25

## Type of Search

NA Sequence (#) 2

AA Sequence (#) 17

Structure (#) \_\_\_\_\_

Bibliographic \_\_\_\_\_

Litigation \_\_\_\_\_

Fulltext \_\_\_\_\_

Patent Family \_\_\_\_\_

Other \_\_\_\_\_

## Vendors and cost where applicable

STN \_\_\_\_\_  
 Dialog \_\_\_\_\_  
 Questia \_\_\_\_\_  
 Lexis/Nexis \_\_\_\_\_  
 CompuServe \_\_\_\_\_  
 Web/Internet \_\_\_\_\_  
 Other (specify) \_\_\_\_\_

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2001, 08:25:09 ; Search time 78.6 Seconds  
(without alignments)  
2485.756 Million cell updates/sec

Title: US-08-887-977-9

Perfect score: 1119

Sequence: 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 302621 seqs, 87301344 residues

Word size : 15

Total number of hits satisfying chosen parameters: 195

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	1.7	3293	1	US-08-030-096-1
2	17	1.5	1032	1	US-08-279-058B-6
3	17	1.5	1071	4	US-09-087-232A-14
4	17	1.5	1344	4	US-09-087-232A-16
5	17	1.5	1376	4	US-09-087-232A-12
6	17	1.5	1414	3	US-08-466-343D-1
7	17	1.5	3791	4	US-08-675-773B-3
8	17	1.5	24979	2	US-08-147-777-3
9	17	1.5	24979	3	US-08-452-872-3
10	17	1.5	24979	5	PCT-US93-03985-3
11	16	1.4	60	1	US-08-089-910-5
12	16	1.4	892	3	US-08-368-768B-1
13	16	1.4	2638	1	US-08-841-349-15
14	16	1.4	1212	2	US-08-449-933-5
15	16	1.4	1220	4	US-09-116-032-2
16	16	1.4	1654	3	US-08-991-426-3
17	16	1.4	1654	3	US-09-143-470-3
18	16	1.4	2097	3	US-09-002-567B-2
19	16	1.4	2638	1	US-08-306-691B-46
20	16	1.4	7400	4	US-09-116-032-1
21	16	1.4	8561	4	US-09-112-450-3
22	16	1.4	8937	2	US-08-449-933-1
23	16	1.4	10706	1	US-08-411-389-1
24	16	1.4	22846	2	US-08-469-461-3
25	16	1.4	22846	3	US-07-890-609-3
26	16	1.4	35100	1	US-08-306-691B-19
27	16	1.4	35100	5	PCT-US93-06251-19

28	16	1.4	246240	2	US-08-724-394A-20	Sequence 20, Appl
29	16	1.4	246240	2	US-08-724-394A-21	Sequence 21, Appl
30	16	1.4	246240	2	US-08-724-394A-22	Sequence 22, Appl
31	15	1.3	32	4	US-09-187-793-7	Sequence 7, Appl
32	15	1.3	48	1	US-08-248-016-2	Sequence 2, Appl
33	15	1.3	48	1	US-08-451-501-2	Sequence 2, Appl
34	15	1.3	48	2	US-08-713-455A-1	Sequence 1, Appl
35	15	1.3	48	5	PCT-US95-06761-2	Sequence 1, Appl
36	15	1.3	54	4	US-09-187-793-8	Sequence 8, Appl
37	15	1.3	127	1	US-08-248-016-8	Sequence 8, Appl
38	15	1.3	127	1	US-08-451-501-8	Sequence 8, Appl
39	15	1.3	127	5	PCT-US95-06761-8	Sequence 8, Appl
40	15	1.3	127	5	PCT-US95-06761-9	Sequence 9, Appl
41	15	1.3	133	1	US-08-248-016-9	Sequence 9, Appl
42	15	1.3	133	1	US-08-451-501-9	Sequence 9, Appl
43	15	1.3	133	2	US-08-713-455A-7	Sequence 7, Appl
44	15	1.3	133	5	PCT-US95-06761-9	Sequence 9, Appl
45	15	1.3	201	2	US-08-592-383-6	Sequence 6, Appl
46	15	1.3	340	3	US-08-441-971-10	Sequence 10, Appl
47	15	1.3	340	4	US-08-836-075A-63	Sequence 63, Appl
48	15	1.3	340	4	US-08-221-653-10	Sequence 10, Appl
49	15	1.3	350	1	US-08-248-016-11	Sequence 11, Appl
50	15	1.3	350	1	US-08-451-501-11	Sequence 11, Appl
51	15	1.3	350	1	PCT-US95-031-173	Sequence 173, Appl
52	15	1.3	350	5	PCT-US95-06761-11	Sequence 11, Appl
53	15	1.3	448	2	US-08-967-101-66	Sequence 66, Appl
54	15	1.3	448	2	US-08-592-541-66	Sequence 66, Appl
55	15	1.3	448	4	US-09-124-698-66	Sequence 66, Appl
56	15	1.3	448	4	US-09-127-480-66	Sequence 66, Appl
57	15	1.3	448	4	US-08-496-841C-66	Sequence 66, Appl
58	15	1.3	463	3	US-08-740-784-2	Sequence 2, Appl
59	15	1.3	529	3	US-08-343-101A-7	Sequence 7, Appl
60	15	1.3	529	3	US-09-183-688-7	Sequence 7, Appl
61	15	1.3	591	1	US-08-725-531-2	Sequence 2, Appl
62	15	1.3	591	2	US-09-213-392-2	Sequence 2, Appl
63	15	1.3	591	2	US-09-083-661-2	Sequence 2, Appl
64	15	1.3	653	1	US-08-181-271A-103	Sequence 103, Appl
65	15	1.3	653	1	US-08-449-315-103	Sequence 103, Appl
66	15	1.3	653	1	US-08-444-803-103	Sequence 103, Appl
67	15	1.3	653	1	US-08-449-043-103	Sequence 103, Appl
68	15	1.3	653	1	US-08-456-265A-103	Sequence 103, Appl
69	15	1.3	653	1	US-08-455-416-103	Sequence 103, Appl
70	15	1.3	653	1	US-08-454-876-103	Sequence 103, Appl
71	15	1.3	653	1	US-08-457-364-103	Sequence 103, Appl
72	15	1.3	653	2	US-08-456-262-103	Sequence 103, Appl
73	15	1.3	653	2	US-08-456-240-103	Sequence 103, Appl
74	15	1.3	653	2	US-08-455-736-103	Sequence 103, Appl
75	15	1.3	653	2	US-08-971-217-103	Sequence 103, Appl
76	15	1.3	653	2	US-08-875-034A-1	Sequence 1, Appl
77	15	1.3	933	3	US-08-808-148-2	Sequence 2, Appl
78	15	1.3	933	3	US-09-134-248-4	Sequence 4, Appl
79	15	1.3	1026	4	US-08-484-993B-21	Sequence 21, Appl
80	15	1.3	1029	2	US-08-484-158B-21	Sequence 21, Appl
81	15	1.3	1029	2	US-08-484-596A-21	Sequence 21, Appl
82	15	1.3	1029	2	US-08-480-150A-21	Sequence 21, Appl
83	15	1.3	1029	3	US-08-458-731-21	Sequence 21, Appl
84	15	1.3	1029	3	US-08-149-223A-21	Sequence 21, Appl
85	15	1.3	1029	3	US-08-454-196-16	Sequence 16, Appl
86	15	1.3	1079	1	US-08-525-503A-3	Sequence 3, Appl
87	15	1.3	1093	3	US-08-196-350-2	Sequence 2, Appl
88	15	1.3	1100	1	US-08-746-883-3	Sequence 3, Appl
89	15	1.3	1101	3	US-08-343-101A-8	Sequence 8, Appl
90	15	1.3	1315	2	US-09-183-688-8	Sequence 8, Appl
91	15	1.3	1315	3	US-09-205-860-1	Sequence 1, Appl
92	15	1.3	1315	3	US-08-861-512-2	Sequence 2, Appl
93	15	1.3	1402	2	US-08-660-347-1	Sequence 1, Appl
94	15	1.3	1458	2	US-07-938-154-10	Sequence 10, Appl
95	15	1.3	1512	2	PCT-US91-02311-10	Sequence 10, Appl
96	15	1.3	1512	5	US-08-865-311-1	Sequence 1, Appl
97	15	1.3	1542	2	US-08-524-828-2	Sequence 2, Appl
98	15	1.3	1593	2	US-08-975-114A-2	Sequence 2, Appl
99	15	1.3	1593	2		
100	15	1.3	1593	2		



RESULT 3  
US-09-087-232A-14  
; Sequence 14, Application, US-09087232A  
; Patent No. 6153431  
; GENERAL INFORMATION:  
; APPLICANT: Quillent et al.  
; TITLE OF INVENTION: HUMAN IMMUNOD

RESULT 3  
US-09-087-232A-14  
; Sequence 14, Application, US-09087232A  
; Patent No. 6153431  
; GENERAL INFORMATION:  
; APPLICANT: Quillent et al.  
; TITLE OF INVENTION: HUMAN IMMUNOD

REFLECTANT: CURRENT ECU:  
 TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR  
 TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole  
 STREET: 30 Rockefeller Plaza  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10112  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/087,232A

REFLECTANT: CURRENT ECU:  
 TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR  
 TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole  
 STREET: 30 Rockefeller Plaza  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10112  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/087,232A

;; FILING DATE: 28 MAY 1998  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/048,057  
;; FILING DATE: 30 MAY 1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: KOLE, LISA B.  
;; REGISTRATION NUMBER: 35,225  
;; REFERENCE/DOCKET NUMBER: AP 31115  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 408-2628  
;; TELEFAX: (212) 765-2519  
;; INFORMATION FOR SEQ ID NO: 16:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1344 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 240..887  
US-09-087-232A-16

Query Match 1.5%; Score 17; DB 4; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951  
|||||  
DB 1112 AGAAGTTCAGAACTAC 1128

RESULT 5  
US-09-087-232A-12  
; Sequence 12, Application US/09087232A  
; Patent No. 6153431  
; GENERAL INFORMATION:  
; APPLICANT: Quillient et al.  
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR  
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/087,232A  
; FILING DATE: 28 MAY 1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/048,057  
; FILING DATE: 30 MAY 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOLE, LISA B.  
; REGISTRATION NUMBER: 35,225  
; REFERENCE/DOCKET NUMBER: AP 31115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 408-2628  
; TELEFAX: (212) 765-2519  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1376 base pairs  
; TYPE: nucleic acid

;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 240..1298  
US-09-087-232A-12

Query Match 1.5%; Score 17; DB 4; Length 1376;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951  
|||||  
DB 1144 AGAAGTTCAGAACTAC 1160

RESULT 6  
US-08-466-343D-1  
; Sequence 1, Application US/08466343D  
; Patent No. 6025154  
; GENERAL INFORMATION:  
; APPLICANT: LI, YI  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN  
; TITLE OF INVENTION: CHEMOKINE RECEPTOR HDGNR10 (AS AMENDED)  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVE., NW, SUITE 500  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,343D  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEFFE, ERIC K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1414 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 259..1314  
US-08-466-343D-1

Query Match 1.5%; Score 17; DB 3; Length 1414;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951  
|||||  
DB 1163 AGAAGTTCAGAACTAC 1179

RESULT 7

US-08-675-773B-3/c  
; Sequence 3, Application US/08675773B  
; Patent No. 6166288  
; GENERAL INFORMATION:  
; APPLICANT: DIAMOND, LISA E  
; APPLICANT: LOGAN, JOHN S  
; APPLICANT: BYRNE, GUERARD W  
; APPLICANT: SHARMA, AJAY  
; TITLE OF INVENTION: METHOD OF PRODUCING TRANSGENIC ANIMALS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
; STREET: 419 SEVENTH STREET, N.W., SUITE 300  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/675,773B  
; FILING DATE: 03-JUL-1996  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOPER, IVER P.  
; REGISTRATION NUMBER: 28,005  
; REFERENCE/DOCKET NUMBER: DIAMOND-1A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3791 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-675-773B-3

Query Match 1.5% Score 17; DB 4; Length 3791;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 901 TGCCTGAACCCCTGTGCT 917  
|||||  
DB 1796 TGCCTGAACCCCTGTGCT 1780

RESULT 8  
US-08-147-777-3/c  
; Sequence 3, Application US/08147777  
; Patent No. 5914265  
; GENERAL INFORMATION:  
; APPLICANT: Roop, Dennis R.  
; APPLICANT: Rothnagel, Joseph A.  
; APPLICANT: Greenhalgh, David A.  
; APPLICANT: Yuspa, Stuart H.  
; TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS  
; TITLE OF INVENTION: AND METHODS OF USE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LYON & LYON  
; STREET: 611 West Sixth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90017  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/147,777  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA: including application  
; PRIOR APPLICATION DATA: described below:  
; APPLICATION NUMBER: 07/876,289  
; FILING DATE: April 30, 1992  
; APPLICATION NUMBER: Unassigned (204/144)  
; FILING DATE: October 29, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 204/153  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 24979 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-147-777-3

Query Match 1.5% Score 17; DB 2; Length 24979;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 878 AAGTCTGGCTTCTCTG 894  
|||||  
DB 502 AAGTCTGGCTTCTCTG 486

RESULT 9  
US-08-452-872-3/c  
; Sequence 3, Application US/08452872  
; Patent No. 6057298  
; GENERAL INFORMATION:  
; APPLICANT: Roop, Dennis R.  
; APPLICANT: Rothnagel, Joseph A.  
; APPLICANT: Greenhalgh, David A.  
; APPLICANT: Yuspa, Stuart H.  
; TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS  
; TITLE OF INVENTION: AND METHODS OF USE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LYON & LYON  
; STREET: 611 West Sixth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/452,872  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/147,777  
; FILING DATE:

APPLICATION NUMBER: 07/876,289  
FILING DATE: April 30, 1992  
APPLICATION NUMBER: Unassigned (204/144)  
FILING DATE: October 29, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 204/153  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24979 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-452-872-3

Query Match 1.5% Score 17; DB 3; Length 24979;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 878 AAGTCCTGGCTTCTCTG 894  
|||||  
DB 502 AAGTCCTGGCTTCTCTG 486

## RESULT 10

PCT-US93-03985-3/c  
Sequence 3, Application PC/TUS9303985  
GENERAL INFORMATION:  
APPLICANT: Roop, Dennis R.  
APPLICANT: Rottnagel, Joseph A.  
APPLICANT: Greenhalgh, David A.  
APPLICANT: Yuspa, Stuart H.  
TITLE OF INVENTION: DEVELOPMENT OF A VECTOR TO TARGET GENE  
EXPRESSION TO THE EPIDERMIS OF TRANSGENIC ANIMALS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski  
STREET: 1301 McKinney, Suite 5100  
CITY: Houston  
STATE: Texas  
COUNTRY: U.S.A.  
ZIP: 77010-3095  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/03985  
FILING DATE: 19930428  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul, Thomas D.  
REGISTRATION NUMBER: 32,714  
REFERENCE/DOCKET NUMBER: D-5478  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713/651-5325  
TELEFAX: 713/651-5246  
TELEX: 762829  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24979 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US93-03985-3

Query Match 1.5% Score 17; DB 5; Length 24979;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 878 AAGTCCTGGCTTCTCTG 894  
|||||  
DB 502 AAGTCCTGGCTTCTCTG 486

## RESULT 11

US-08-089-910-5  
Sequence 5, Application US/08089910  
Patent No. 5460941  
GENERAL INFORMATION:  
APPLICANT: Camerini-Otero, Raphael D.  
APPLICANT: McIntosh, Margaret  
APPLICANT: Camerini-Otero, Carol S.  
APPLICANT: Ferrin, Lance J.  
TITLE OF INVENTION: METHOD OF TARGETING DNA  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/089,910  
FILING DATE: 12-JUL-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fuller, Michael L.  
REGISTRATION NUMBER: 36,516  
REFERENCE/DOCKET NUMBER: NIH068.001FW1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 60 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-089-910-5

Query Match 1.4% Score 16; DB 1; Length 60;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1029 CTCAGAAACATTCT 1044  
|||||  
DB 7 CTCAGAAACATTCT 22

RESULT 12  
US-08-968-768B-1/c  
; Sequence 1, Application US/08968768B  
; Patent No. 6083719  
; GENERAL INFORMATION:  
; APPLICANT: Momparler, Richard L  
; APPLICANT: Laliberte, Josee  
; APPLICANT: Eliopoulos, Nicoletta  
; APPLICANT: Cournoyer, Denis  
; TITLE OF INVENTION: Cytidine deaminase cDNA as a positive  
; TITLE OF INVENTION: selectable marker for gene transfer, gene therapy and for  
; TITLE OF INVENTION: protein synthesis  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Swabey Ogilvy Renault  
; STREET: 1981 McGill College, suite 1600  
; CITY: Montreal  
; STATE: Quebec  
; COUNTRY: Canada  
; ZIP: H3A 2X3  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/968,768B  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/509,138  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cote, France  
; REGISTRATION NUMBER: 37,037  
; REFERENCE/DOCKET NUMBER: 12667-3US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 514 845 7126  
; TELEFAX: 514 288 8389  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 892 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; TISSUE TYPE: Myeloid  
; CELL LINE: HL-60  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT: 1  
; FEATURE:  
; NAME: CDS  
; LOCATION: 118..558  
; PUBLICATION INFORMATION:  
; AUTHORS: Laliberte, Josee  
; AUTHORS: Momparler, Richard L  
; TITLE: Human Cytidine Deaminase: Purification of  
; TITLE: Enzyme, Cloning, and Expression of its  
; TITLE: complementary DNA  
; JOURNAL: Cancer Research  
; VOLUME: 54  
; PAGES: 5401-5407  
; DATE: October 15-1994  
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 922

Query Match 1.4%; Score 16; DB 3; Length 892;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 146 TTCCTGGCCTCTCTGGG 161  
Db 197 TTCCTGGCCTCTCTGGG 182  
  
RESULT 13  
US-08-841-349-15/c  
; Sequence 15, Application US/08841349B  
; Patent No. 5955594  
; GENERAL INFORMATION:  
; APPLICANT: MISHRA, LOPA  
; TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT.  
; FILE REFERENCE: XX/P044700SU  
; CURRENT APPLICATION NUMBER: US/08/841,349B  
; CURRENT FILING DATE: 1997-04-30  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 981  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: For all n's in this sequence, n=(a or g or c or t)  
US-08-841-349-15  
  
Query Match 1.4%; Score 16; DB 2; Length 981;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 96 AGGCAGTTCTCCAGGC 111  
Db 788 AGGCAGTTCTCCAGGC 773  
  
RESULT 14  
US-08-449-933-5  
; Sequence 5, Application US/08449933  
; Patent No. 5859195  
; GENERAL INFORMATION:  
; APPLICANT: Collins, Francis S.  
; APPLICANT: Wallace, Margaret R.  
; APPLICANT: Marchuk, Douglas A.  
; APPLICANT: Anderson, Lone B.  
; APPLICANT: Guttman, David H.  
; TITLE OF INVENTION: Neurofibromatosis Gene  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/449,933  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Konski, Antoinette F.  
; REGISTRATION NUMBER: 34,202  
; REFERENCE/DOCKET NUMBER: 20344-20553.10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792  
TELEX: 706141 MRSNFOERS SFO  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1212 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 211..1212  
NAME/KEY: misc\_feature  
LOCATION: 52..54  
OTHER INFORMATION: /note= "Upstream in frame stop"  
OTHER INFORMATION: codon  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 98..119  
OTHER INFORMATION: /note= "Oligonucleotide used for primer extension"  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (270~271)  
OTHER INFORMATION: /note= "Position of the first intron and alternate sequences (SEQ ID NO: 6 through SEQ ID NO: 7)"  
OTHER INFORMATION: diverge"  
US-08-449-933-5

Query Match 1.4%; Score 16; DB 2; Length 1212;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 TTGATGTCATGATAT 667  
|||||  
DB 725 TTGATGTCATGATAT 740

RESULT 15  
US-09-116-032-2/c  
Sequence 2, Application US/09116032  
Patent No. 6200576  
GENERAL INFORMATION:  
APPLICANT: HWONG, CHING LONG  
APPLICANT: LO, CHENG-KAI  
APPLICANT: YANG, YING-CHUAN  
APPLICANT: JENG, KING-SONG  
APPLICANT: CHANG, EDWARD L.  
APPLICANT: DEVELOPMENT CENTER FOR BIOTECHNOLOGY  
TITLE OF INVENTION: SWINE VESICULAR DISEASE VIRUS AND MUTANT STRAINS AND PREPARATION PROCESS AND USE THEREOF  
FILE REFERENCE: 9751.79US01  
CURRENT APPLICATION NUMBER: US/09/116,032  
CURRENT FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: CHINA 86105814  
EARLIER FILING DATE: 1997-05-01  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1220  
TYPE: DNA  
ORGANISM: SWINE VESICULAR DISEASE VIRUS  
US-09-116-032-2

Query Match 1.4%; Score 16; DB 4; Length 1220;  
Best Local Similarity 100.0%; Pred. No. 78;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 293 CCACTGGTGGTGGT 308  
|||||  
DB 505 CCACTGGTGGTGGT 490

RESULT 16  
US-08-991-426-3/c  
Sequence 3, Application US/08991426  
Patent No. 6013257  
GENERAL INFORMATION:  
APPLICANT: Pan, Yang  
TITLE OF INVENTION: NEUROTACTIN AND USES THEREFOR  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,426  
FILING DATE: 16-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/851,160  
FILING DATE: 05-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/643,798  
FILING DATE: 07-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meiklejohn, Ph.D., Anita L.  
REGISTRATION NUMBER: 35,283  
REFERENCE/DOCKET NUMBER: 09404/043001  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1654 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 86...1276  
US-08-991-426-3

Query Match 1.4%; Score 16; DB 3; Length 1654;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 TGGCCTCCTGGGGAAT 165  
|||||  
DB 1448 TGGCCTCCTGGGGAAT 1433

RESULT 17  
US-09-143-470-3/c  
Sequence 3, Application US/09143470  
Patent No. 6043086  
GENERAL INFORMATION:  
APPLICANT: Pan, Yang  
TITLE OF INVENTION: NEUROTACTIN AND USES THEREFOR

FILE REFERENCE: 09404/049001  
CURRENT APPLICATION NUMBER: US/09/143,470  
CURRENT FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 1654  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (86)...(1276)  
US-09-143-470-3

Query Match 1.4%; Score 16; DB 3; Length 1654;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 150 TGGCTCTCTGGGAAT 165  
|||||  
Db 1448 TGGCTCTCTGGGAAT 1433

## RESULT 18

US-09-002-567B-2  
Sequence 2, Application US/09002567B  
Patent No. 6001594  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN TESTIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/002,567B  
FILING DATE: December 31, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0454 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2097 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: NEUTEMT01  
CLONE: 338680  
US-09-002-567B-2

Query Match 1.4%; Score 16; DB 3; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 78;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 928 ATTGGGCAGAGTTCA 943  
|||||  
Db 1343 ATTGGGCAGAGTTCA 1358

## RESULT 19

US-08-306-691B-46/C  
Sequence 46, Application US/08306691B  
Patent No. 5734039  
GENERAL INFORMATION:  
APPLICANT: Calabretta, Bruno  
APPLICANT: Skorski, Tomasz  
TITLE OF INVENTION: ANTISENSE  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seidel, Gonda, Lavoigna & Monaco, P.C.  
STREET: Two Penn Center, Suite 1800  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19102

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/306,691B  
FILING DATE: September 15, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8321-8  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5734039e  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2638 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-306-691B-46

Query Match 1.4%; Score 16; DB 1; Length 2638;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 934 CAGAAAGTTCAGAAACT 949  
|||||  
Db 1506 CAGAAAGTTCAGAAACT 1491

## RESULT 20

US-09-116-032-1/c  
Sequence 1, Application US/09116032  
Patent No. 6200576  
GENERAL INFORMATION:  
APPLICANT: HWONG, CHING LONG  
APPLICANT: LO, CHENG-KAI  
APPLICANT: YANG, YING-CHUAN  
APPLICANT: JENG, KING-SONG  
APPLICANT: CHANG, EDWARD L.  
APPLICANT: DEVELOPMENT CENTER FOR BIOTECHNOLOGY

```
; TITLE OF INVENTION: SWINE VESICULAR DISEASE VIRUS AND MUTANT STRAINS AND
; FILE OF INVENTION: PREPARATION PROCESS AND USE THEREOF
; FILE REFERENCE: 9751.79US01
; CURRENT APPLICATION NUMBER: US/09/116,032
; CURRENT FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: CHINA 86105814
; EARLIER FILING DATE: 1997-05-01
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 7400
; TYPE: DNA
; ORGANISM: SWINE VESICULAR DISEASE VIRUS
; US-09-116-032-1

Query Match      1.4%; Score 16; DB 4; Length 7400;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 CCACTGTCGTCGGGT 308
      |||||
Db 505 CCACTGTCGTCGGGT 490

RESULT 21
US-09-112-450-3/c
; Sequence 3, Application US/09112450
; Patent No. 6120999
; GENERAL INFORMATION:
; APPLICANT: Abad, Antonio Jose C.
; APPLICANT: Choi, Gil
; APPLICANT: Calderone, Richard A.
; TITLE OF INVENTION: Histidine Kinase Two-component in Candida albicans
; FILE REFERENCE: PF393
; CURRENT APPLICATION NUMBER: US/09/112,450
; CURRENT FILING DATE: 1998-07-09
; EARLIER APPLICATION NUMBER: 60/052,273
; EARLIER FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: 60/074,308
; EARLIER FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 8561
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-112-450-3

Query Match      1.4%; Score 16; DB 4; Length 8561;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 682 ATTGTCAAACCTTGG 697
      |||||
Db 1479 ATTGTCAAACCTTGG 1464

RESULT 22
US-08-449-933-1
; Sequence 1, Application US/08449933
; Patent No. 5859195
; GENERAL INFORMATION:
; APPLICANT: Collins, Francis S.
; APPLICANT: Wallace, Margaret R.
; APPLICANT: Marchuk, Douglas A.
; APPLICANT: Anderson, Lone B.
; APPLICANT: Guttman, David H.
; TITLE OF INVENTION: Neurofibromatosis Gene
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
```

```
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,933
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Konski, Antoinette F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: 20344-20553.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141 MRSNFOERS SFO
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8937 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 17q11.2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6814..8937
; OTHER INFORMATION: /note= "Entire length of sequence"
; OTHER INFORMATION: clone p5"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..8646
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8425..8646
; OTHER INFORMATION: /note= "219 nt PstI-HindIII"
; OTHER INFORMATION: fragment designated pMAL.B3A"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382..1302
; OTHER INFORMATION: /note= "918 nt HpaI-PstI fragment"
; OTHER INFORMATION: designated pMAL.HF3A.P"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382..3909
; OTHER INFORMATION: /note= "3523 nt HpaI-XhoI fragment"
; OTHER INFORMATION: designated pMAL.HF3A.X"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8383..8937
; OTHER INFORMATION: /note= "Clone B3A"
; PUBLICATION INFORMATION:
; AUTHORS: Wallace, M.R. et al.
; TITLE: Type 1 Neurofibromatosis Gene: Correction
; JOURNAL: Science
; VOLUME: 250
; ISSUE: 12/21/90
; PAGES: 1749-
; DATE: 12/21-1990
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 8937
; PUBLICATION INFORMATION:
; AUTHORS: Wallace, M.R. et al.
```

;  
; TITLE: Type 1 Neurofibromatosis Gene: Identification  
; TITLE: of a Large Transcript in Three NF1 Patients  
; JOURNAL: Science  
; VOLUME: 249  
; ISSUE: 07/13/90  
; PAGES: 181-186  
; DATE: 07/13-1990  
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 8937  
US-08-449-933-1

Query Match 1.4%; Score 16; DB 2; Length 8937;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 TTGATGTTTCATGATAT 667  
|||||  
DB 704 TTGATGTTTCATGATAT 719

RESULT 23  
US-08-411-389-1

; Sequence 1, Application US/08411389  
; Patent No. 5605799  
; GENERAL INFORMATION:  
; APPLICANT: White, Raymond L.  
; APPLICANT: Cawthon, Richard M.  
; APPLICANT: Li, Ying

; TITLE OF INVENTION: SOMATIC MUTATIONS IN THE  
; TITLE OF INVENTION: NEUROFIBROMATOSIS TYPE 1 GENE IN HUMAN TUMORS  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
; STREET: 1201 New York Avenue NW, Suite 1000  
; CITY: Washington  
; STATE: DC  
; ZIP: 20005

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/411,389  
; FILING DATE:  
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/047,088  
; FILING DATE: 16-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ihnen, Jeffrey L.  
; REGISTRATION NUMBER: 28,957  
; REFERENCE/DOCKET NUMBER: 19780-107116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-962-4810  
; TELEFAX: 202-962-8300

; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10706 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 287...8740  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 3809..4888

; OTHER INFORMATION: /function= "NF1 GRD"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 4451..4459  
; OTHER INFORMATION: /function= "Conserved area in GRD"  
US-08-411-389-1

Query Match 1.4%; Score 16; DB 1; Length 10706;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 TTGATGTTTCATGATAT 667  
|||||  
DB 801 TTGATGTTTCATGATAT 816

RESULT 24

US-08-469-461-3  
; Sequence 3, Application US/08469461B  
; Patent No. 5981178  
; GENERAL INFORMATION:

; APPLICANT: Tsui, Lap-Chee  
; APPLICANT: Rommins, Johanna M.  
; APPLICANT: Kerem, Bat-Sheva  
; TITLE OF INVENTION: Introns and Exons of the Cystic Fibrosis Gene and  
; TITLE OF INVENTION: Mutations at Various Positions of the Gene  
; FILE REFERENCE: 3477-61, 033477/139840  
; CURRENT APPLICATION NUMBER: US/08/469,461B  
; CURRENT FILING DATE: 1995-06-06  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 22846  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-469-461-3

Query Match 1.4%; Score 16; DB 2; Length 22846;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1029 CTCAGAAAACATTCT 1044  
|||||  
DB 1294 CTCAGAAAACATTCT 1309

RESULT 25

US-07-890-609-3  
; Sequence 3, Application US/07890609C  
; Patent No. 6001588  
; GENERAL INFORMATION:

; APPLICANT: Tsui, Lap-Chee  
; APPLICANT: Rommins, Johanna M.  
; APPLICANT: Kerem, Bat-Sheva  
; TITLE OF INVENTION: Introns and Exons of the Cystic Fibrosis Gene and  
; TITLE OF INVENTION: Mutations at Various Positions of the Gene  
; FILE REFERENCE: 3477-61, 033477/139840  
; CURRENT APPLICATION NUMBER: US/07/890,609C  
; CURRENT FILING DATE: 1992-07-13  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 22846  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-07-890-609-3

Query Match 1.4%; Score 16; DB 3; Length 22846;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1029 CTCAGAAACATTCT 1044  
|||||  
Db 1294 CTCAGAAACATTCT 1309

## RESULT 26

US-08-306-691B-19/c  
; Sequence 19, Application US/08306691B  
; Patent No. 5734039  
; GENERAL INFORMATION:  
; APPLICANT: Calabretta, Bruno  
; APPLICANT: Skorski, Tomas  
; TITLE OF INVENTION: ANTISENSE  
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.  
; STREET: Two Penn Center, Suite 1800  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306.691B  
; FILING DATE: September 15, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8321-8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; TELEX: No. 5734039e  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35100 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-306-691B-19

Query Match 1.4%; Score 16; DB 1; Length 35100;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 TCTGCCCTTGTGTGTG 488  
|||||  
Db 26288 TCTGCCCTTGTGTGTG 26273

## RESULT 27

PCT-US93-06251-19/c  
; Sequence 19, Application PC/TUS9306251  
; GENERAL INFORMATION:  
; APPLICANT: Wickstrom, Eric and Rife, Jason P.  
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City

; STATE: NY  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/06251  
; FILING DATE: 19930630  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8586  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35100 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
PCT-US93-06251-19

Query Match 1.4%; Score 16; DB 5; Length 35100;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 TCTGCCCTTGTGTGTG 488  
|||||  
Db 26288 TCTGCCCTTGTGTGTG 26273

## RESULT 28

US-08-724-394A-20  
; Sequence 20, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Kronmal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchinashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1  
; TITLE OF INVENTION: Sequences and Antibodies Thereto  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,394A  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136

REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-20

Query Match 1.4%; Score 16; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 ATGATATTTGTACA 676  
|||||  
DB 38333 ATGATATTTGTACA 38348

RESULT 29  
US-08-724-394A-21  
Sequence 21, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Kronmal, Gregory S.  
APPLICANT: Lauer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereto  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:

NAME/KEY: misc\_feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-21

Query Match 1.4%; Score 16; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 ATGATATTTGTACA 676  
|||||  
DB 38333 ATGATATTTGTACA 38348

RESULT 30  
US-08-724-394A-22  
Sequence 22, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Kronmal, Gregory S.  
APPLICANT: Lauer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereto  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-22

Query Match 1.4%; Score 16; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 661 ATGATATTTGTACA 676  
|||||

```
Db 38333 ATCATATTTTGTAC 38348

RESULT 31
US-09-187-793-7
; Sequence 7, Application US/09187793C
; Patent No. 6136537
; GENERAL INFORMATION:
; APPLICANT: Macevitz, Stephen C.
; TITLE OF INVENTION: Gene Expression Analysis
; FILE REFERENCE: 1002-01
; CURRENT APPLICATION NUMBER: US/09/187,793C
; CURRENT FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: US 09/028,128
; PRIOR FILING DATE: 1998-02-23
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Microsoft Word 2.x for Windows
; SEQ ID NO 7
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION: Single strand of adaptor
US-09-187-793-7

Query Match 1.3%; Score 15; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 889 TTCCTGCACCTGCTGC 903
|||||
Db 8 ttctgcactgctgc 22

RESULT 32
US-08-248-016-2/c
; Sequence 2, Application US/08248016
; Patent No. 5550109
; GENERAL INFORMATION:
; APPLICANT: Schonwetter, Barry S.
; APPLICANT: Zasloff, Michael A.
; TITLE OF INVENTION: Inducible Defensin Peptide From
; TITLE OF INVENTION: Mammalian Epithelia
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,016
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 30,907
; REFERENCE/DOCKET NUMBER: 05387.0017-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,016
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ogden, Stasia L.
; REGISTRATION NUMBER: 36,228
; REFERENCE/DOCKET NUMBER: 05387.0017-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-248-016-2

Query Match 1.3%; Score 15; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGTCTCGGAGCCCA 595
|||||
Db 34 CTGTCTCGGAGCCCA 20

RESULT 33
US-08-451-501-2/c
; Sequence 2, Application US/08451501
; Patent No. 5656738
; GENERAL INFORMATION:
; APPLICANT: Schonwetter, Barry S.
; APPLICANT: Zasloff, Michael A.
; TITLE OF INVENTION: Inducible Defensin Peptide From
; TITLE OF INVENTION: Mammalian Epithelia
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,501
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,016
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 30,907
; REFERENCE/DOCKET NUMBER: 05387.0017-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-451-501-2

Query Match 1.3%; Score 15; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGTCTCGGAGCCCA 595
|||||
Db 34 CTGTCTCGGAGCCCA 20
```

RESULT 34  
US-08-713-455A-1/c  
Sequence 1, Application US/08713455A  
Patent No. 5849490  
GENERAL INFORMATION:  
APPLICANT: Schonwetter, Barry S.  
ATTORNEY/AGENT INFORMATION:  
NAME: Zasloff, Michael A.  
TITLE OF INVENTION: INDUCIBLE DEFENSIN PEPTIDE FROM  
TITLE OF INVENTION: MAMMALIAN EPITHELIA  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner, LLP  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/713,455A  
FILING DATE: 13-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lonigan, Simone L.  
REFERENCE/DOCKET NUMBER: 05387.0017-02000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)408-4366  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-08-713-455A-1

Query Match 1.3%; Score 15; DB 2; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGCTCGGAGCCCA 595  
|||||  
Db 34 CTGCTCGGAGCCCA 20

RESULT 35  
PCT-US95-06761-2/c  
Sequence 2, Application PC/TUS9506761  
GENERAL INFORMATION:  
APPLICANT: Magainin Pharmaceuticals Inc.  
APPLICANT: 5110 Campus Drive  
APPLICANT: Plymouth Meeting, PA 19462  
TITLE OF INVENTION: Inducible Defensin Peptide From  
TITLE OF INVENTION: Mammalian Epithelia  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06761  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,016  
FILING DATE: 24-MAY-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ogden, Stasia L.  
REGISTRATION NUMBER: 36,228  
REFERENCE/DOCKET NUMBER: 05387.0017-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
PCT-US95-06761-2

Query Match 1.3%; Score 15; DB 5; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGCTCGGAGCCCA 595  
|||||  
Db 34 CTGCTCGGAGCCCA 20

RESULT 36  
US-09-187-793-8/c  
Sequence 8, Application US/09187793C  
Patent No. 6136537  
GENERAL INFORMATION:  
APPLICANT: Macevitz, Stephen C.  
TITLE OF INVENTION: Gene Expression Analysis  
FILE REFERENCE: 1002-01  
CURRENT APPLICATION NUMBER: US/09/187,793C  
CURRENT FILING DATE: 1998-11-06  
PRIOR APPLICATION NUMBER: US 09/028,128  
PRIOR FILING DATE: 1998-02-23  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Microsoft Word 2.x for Windows  
SEQ ID NO 8  
LENGTH: 54  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY:  
LOCATION:  
OTHER INFORMATION: Double stranded insert  
US-09-187-793-8

Query Match 1.3%; Score 15; DB 4; Length 54;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 889 TTCTGTGCACTGCTGC 903  
|||||  
Db 29 TTCTGTGCACTGCTGC 15

RESULT 37  
US-08-248-016-8  
Sequence 8, Application US/08248016

```
; Patent No. 5550109
; GENERAL INFORMATION:
; APPLICANT: Schonwetter, Barry S.
; APPLICANT: Zasloff, Michael A.
; TITLE OF INVENTION: Inducible Defensin Peptide From
; TITLE OF INVENTION: Mammalian Epithelia
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ATTORNEY/AGENT INFORMATION:
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248.016
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ogden, Stasia L.
; REGISTRATION NUMBER: 36,228
; REFERENCE/DOCKET NUMBER: 05387.0017-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-248-016-8

Query Match 1.3%; Score 15; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGTCGAGCCCA 595
Db 90 CTGTCGAGCCCA 104
|||||

RESULT 38
US-08-451-501-8
; Sequence 8, Application US/08451501
; Patent No. 5556738
; GENERAL INFORMATION:
; APPLICANT: Schonwetter, Barry S.
; APPLICANT: Zasloff, Michael A.
; TITLE OF INVENTION: Inducible Defensin Peptide From
; TITLE OF INVENTION: Mammalian Epithelia
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ATTORNEY/AGENT INFORMATION:
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451.501
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,016
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 30,907
; REFERENCE/DOCKET NUMBER: 05387.0017-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-451-501-8

Query Match 1.3%; Score 15; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGTCGAGCCCA 595
Db 90 CTGTCGAGCCCA 104
|||||

RESULT 39
US-08-713-455A-6
; Sequence 6, Application US/08713455A
; Patent No. 5849490
; GENERAL INFORMATION:
; APPLICANT: Schonwetter, Barry S.
; APPLICANT: Zasloff, Michael A.
; TITLE OF INVENTION: INDUCIBLE DEFENSIN PEPTIDE FROM
; TITLE OF INVENTION: MAMMALIAN EPITHELIA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ATTORNEY/AGENT INFORMATION:
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,455A
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lonigan, Simone L.
; REFERENCE/DOCKET NUMBER: 05387.0017-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4366
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

; MOLECULE TYPE: CDNA  
US-08-713-455A-6

Query Match 1.3%; Score 15; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 581 CTGTCTCGGAGCCCA 595  
Db 90 CTGTCTCGGAGCCCA 104  
|||||

## RESULT 40

PCT-US95-06761-8  
; Sequence 8, Application PC/TUS9506761  
; GENERAL INFORMATION:  
; APPLICANT: Magainin Pharmaceuticals Inc.  
; APPLICANT: 5110 Campus Drive  
; APPLICANT: Plymouth Meeting, PA 19462  
; TITLE OF INVENTION: Inducible Defensin Peptide From  
; TITLE OF INVENTION: Mammalian Epithelia  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/06761  
; FILING DATE: 24-MAY-1994

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/248,016  
; FILING DATE: 24-MAY-1994

; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ogden, Stasia L.  
; REGISTRATION NUMBER: 36,228  
; REFERENCE/DOCKET NUMBER: 05387.0017-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400

; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 127 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
PCT-US95-06761-8

Query Match 1.3%; Score 15; DB 5; Length 127;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 581 CTGTCTCGGAGCCCA 595  
Db 90 CTGTCTCGGAGCCCA 104  
|||||

## RESULT 41

US-08-248-016-9  
; Sequence 9, Application US/08248016  
; Patent No. 5550109

; GENERAL INFORMATION:  
; APPLICANT: Schonwetter, Barry S.  
; APPLICANT: Zasloff, Michael A.  
; TITLE OF INVENTION: Inducible Defensin Peptide From  
; TITLE OF INVENTION: Mammalian Epithelia  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/248,016  
; FILING DATE: 24-MAY-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ogden, Stasia L.  
; REGISTRATION NUMBER: 36,228  
; REFERENCE/DOCKET NUMBER: 05387.0017-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 133 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
US-08-248-016-9

Query Match 1.3%; Score 15; DB 1; Length 133;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 581 CTGTCTCGGAGCCCA 595  
Db 96 CTGTCTCGGAGCCCA 110  
|||||

## RESULT 42

US-08-451-501-9  
; Sequence 9, Application US/08451501  
; Patent No. 5656738

; GENERAL INFORMATION:  
; APPLICANT: Schonwetter, Barry S.  
; APPLICANT: Zasloff, Michael A.  
; TITLE OF INVENTION: Inducible Defensin Peptide From  
; TITLE OF INVENTION: Mammalian Epithelia  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/451,501  
FILING DATE: 26-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,016  
FILING DATE: 24-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 30,907  
REFERENCE/DOCKET NUMBER: 05387.0017-01000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4400  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-451-501-9

Query Match 1.3%; Score 15; DB 1; Length 133;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGCTCGGAGCCCA 595  
DB 96 CTGCTCGGAGCCCA 110

RESULT 43  
US-08-713-455A-7  
Sequence 7, Application US/08/713455A  
Patent No. 5849490  
GENERAL INFORMATION:  
APPLICANT: Schonwetter, Barry S.  
APPLICANT: Zasloff, Michael A.  
TITLE OF INVENTION: INDUCIBLE DEFENSIN PEPTIDE FROM  
TITLE OF INVENTION: MAMMALIAN EPITHELIA  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner, LLP  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/713,455A  
FILING DATE: 13-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lonigan, Simone L.  
REFERENCE/DOCKET NUMBER: 05387.0017-02000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)408-4366  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

US-08-713-455A-7

Query Match 1.3%; Score 15; DB 2; Length 133;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGCTCGGAGCCCA 595  
DB 96 CTGCTCGGAGCCCA 110

RESULT 44  
PCT-US95-06761-9  
Sequence 9, Application PC/TUS9506761  
GENERAL INFORMATION:  
APPLICANT: Magainin Pharmaceuticals Inc.  
APPLICANT: 5110 Campus Drive  
APPLICANT: Plymouth Meeting, PA 19462  
TITLE OF INVENTION: Inducible Defensin Peptide From  
TITLE OF INVENTION: Mammalian Epithelia  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06761  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,016  
FILING DATE: 24-MAY-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ogden, Stasia L.  
REGISTRATION NUMBER: 36,228  
REFERENCE/DOCKET NUMBER: 05387.0017-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
PCT-US95-06761-9

Query Match 1.3%; Score 15; DB 5; Length 133;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 CTGCTCGGAGCCCA 595  
DB 96 CTGCTCGGAGCCCA 110

RESULT 45  
US-08-592-383-6  
Sequence 6, Application US/08592383  
Patent No. 5830760  
GENERAL INFORMATION:

APPLICANT: Tsai, S. and S.J. Collins  
TITLE OF INVENTION: "Hematopoietic Cell Lines Bearing Altered Retinoic Acid Recept  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christensen, O'Connor, Johnson and Kindness  
STREET: 2800 Pacific First Centre, 1420 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage  
COMPUTER: IBM PC/386 Compatible  
OPERATING SYSTEM: MS-DOS 4.01  
SOFTWARE: Word for Windows 5.01-t  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/592,383  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/099,242  
FILING DATE: July 28, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Broderick, Thomas F.  
REGISTRATION NUMBER: 31,332  
REFERENCE/DOCKET NUMBER: FHC-1-7190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-206-682-8100; 1-206-224-0709(direct)  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: page 11, RAR-alpha DNA binding region; positions 364 to 564 of SEQ.  
IMMEDIATE SOURCE:  
LIBRARY: CDNA  
US-08-592-383-6

Query Match 1.3%; Score 15; DB 2; Length 201;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 995 ACAAGTCCTCAGGCT 1009  
Db 17 ACAAGTCCTCAGGCT 31  
RESULT 46  
US-08-441-971-10/c  
Sequence 10, Application US/08441971  
Patent No. 6071693  
GENERAL INFORMATION:  
APPLICANT: Tai-An Cha  
TITLE OF INVENTION: HCV GENOMIC SEQUENCES FOR  
DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS Version 3.3  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/441,971

FILING DATE: 16-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/221,653  
FILING DATE:  
APPLICATION NUMBER: US/07/881,528  
FILING DATE:  
APPLICATION NUMBER: 07/697,326  
FILING DATE: 8 May 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Janluk, Anthony J.  
REGISTRATION NUMBER: 29,809  
REFERENCE/DOCKET NUMBER: C0772/7000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 720-3500  
TELEFAX: (617) 720-2441  
TELEX: EZEKIEL  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: ns5gh6  
US-08-441-971-10

Query Match 1.3%; Score 15; DB 3; Length 340;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1062 CGCAGATAACGACAA 1076  
Db 299 CGCAGATAACGACAA 285

RESULT 47  
US-08-836-075A-63/c  
Sequence 63, Application US/08836075A  
Patent No. 6180768  
GENERAL INFORMATION:  
APPLICANT: MAERTENS, GEERT  
APPLICANT: STUYVER, LIEVEN  
TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES  
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC  
TITLE OF INVENTION: AGENTS  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE & DURKEE  
STREET: P.O. BOX 4433  
CITY: HOUSTON  
STATE: TEXAS  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Microsoft Word 6.0 / ASCII text output  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/836,075A  
FILING DATE: 21 Apr 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP95/04155  
FILING DATE: 23 Oct 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 94870166.9  
FILING DATE: 21 Oct 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95870076.7  
FILING DATE: 28 Jun 1995

## ATTORNEY/AGENT INFORMATION:

NAME: KAMMERER, PATRICIA A.  
REGISTRATION NUMBER: 29,775  
REFERENCE/DOCKET NUMBER: INNS:004  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-836-075A-63

Query Match 1.3%; Score 15; DB 4; Length 340;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1062 CGCAGATAACGACAA 1076  
|||||  
Db 299 CGCAGATAACGACAA 285

## RESULT 48

US-08-221-653-10/c  
; Sequence 10, Application US/08221653  
; Patent No. 6190864

## GENERAL INFORMATION:

APPLICANT: Tai-An Cha  
TITLE OF INVENTION: HCV GENOMIC SEQUENCES FOR  
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02210

COMPUTER READABLE FORM:  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS Version 3.3  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/221,653  
FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/881,528  
FILING DATE:

APPLICATION NUMBER: 07/697,326  
FILING DATE: 8 May 1991

## ATTORNEY/AGENT INFORMATION:

NAME: Janiuk, Anthony J.  
REGISTRATION NUMBER: 29,809  
REFERENCE/DOCKET NUMBER: C0772/7000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 720-3500  
TELEFAX: (617) 720-2441  
TELEX: EZEKIEL

## INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:  
LENGTH: 340 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
ORIGINAL SOURCE: ns5gh6  
INDIVIDUAL ISOLATE: ns5gh6  
US-08-221-653-10

## ATTORNEY/AGENT INFORMATION:

NAME: KAMMERER, PATRICIA A.  
REGISTRATION NUMBER: 29,775  
REFERENCE/DOCKET NUMBER: INNS:004  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-836-075A-63

## RESULT 49

US-08-248-016-11  
; Sequence 11, Application US/08248016  
; Patent No. 5550109

## GENERAL INFORMATION:

APPLICANT: Schonwetter, Barry S.  
APPLICANT: Zasloff, Michael A.  
TITLE OF INVENTION: Inducible Defensin Peptide From  
TITLE OF INVENTION: Mammalian Epithelia  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315

## COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/248,016  
FILING DATE: 24-MAY-1994  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Ogden, Stasia L.  
REGISTRATION NUMBER: 36,228  
REFERENCE/DOCKET NUMBER: 05387.0017-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 350 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-248-016-11

Query Match 1.3%; Score 15; DB 1; Length 350;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 581 CTGTCTCGGAGCCCA 595  
|||||  
Db 179 CTGTCTCGGAGCCCA 193

## RESULT 50

US-08-451-501-11  
; Sequence 11, Application US/08451501  
; Patent No. 5656738

## GENERAL INFORMATION:

APPLICANT: Schonwetter, Barry S.  
APPLICANT: Zasloff, Michael A.  
TITLE OF INVENTION: Inducible Defensin Peptide From  
TITLE OF INVENTION: Mammalian Epithelia  
US-08-221-653-10

NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/451,501  
FILING DATE: 26-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/248,016  
FILING DATE: 24-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 30,907  
REFERENCE/DOCKET NUMBER: 05387.0017-01000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 350 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-08-451-501-11

Query Match 1.38; Score 15; DB 1; Length 350;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 581 CTGCTCGGAGCCCA 595  
|||||  
Db 179 CTGCTCGGAGCCCA 193

Search completed: May 24, 2001, 09:18:17  
Job time: 3188 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2001, 06:53:09 ; Search time 1452.25 Seconds  
(without alignments)  
11363.646 Million cell updates/sec

Title: us-08-887-977-9  
Perfect score: 1119  
Sequence: 1 ATGTTTTCAGTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1283235 seqs, 7373929652 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vi:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_vil:\*  
59: gb\_vil2:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_rol:\*  
95: gb\_rol2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1100.6	98.4	1255	92	HSCKRL3	279784 H.sapiens G
2	1100.6	98.4	3693	93	HSU45984	U45984 Homo sapien
3	1099.	98.2	101230	92	HSBA517H2	AL121935 Human DNA
4	1097.4	98.1	1518	93	HSU68032	U68032 Human G pro
5	1097.4	98.1	2978	93	HSU68030	U68030 Human G pro
6	1065.2	95.2	1137	93	HSU60000	U60000 Human IL8-r
c 7	876.6	78.3	205272	79	AL353591	AL353591 Homo sapi
8	631	56.4	1302	94	AB009369	AB009369 Mus muscu
9	631	56.4	1674	94	MSM222714	AJ222714 Mus muscu
10	624.6	55.8	1410	94	AB016031	AB016031 Mus muscu
11	222	19.8	1500	92	HSDNABLR2	X84702 H.sapiens B

12	222	19.8	1900	9	ARI07241	Sequence	
13	222	19.8	2058	9	ARI07232	Sequence	
14	222	19.8	2139	93	HUMEB1ICDN	L31581 Human G pro	
15	222	19.8	2160	9	ARI07239	Sequence	
16	222	19.8	2215	93	HUMEB1I03	L31584 Human G pro	
17	222	19.8	13801	85	AC004585	Homo sapi	
18	217.2	19.4	106991	65	AC016960	Homo sapi	
c	18	217.2	19.4	106991	65	AC016960	Homo sapi
19	216.2	19.3	2072	94	MUSEB1ICDN	L31580 Mouse G pro	
20	216.2	19.3	2085	9	ARI07278	Sequence	
21	215.6	19.3	2154	9	AR003349	Sequence	
22	215.6	19.3	2154	9	AR008561	Sequence	
23	215.6	19.3	2154	93	HUMGPCRA	L08176 Human Epste	
24	211.8	18.9	1608	94	MMU132336	Mus muscu	
25	211.8	18.9	2843	94	MMU131357	Mus muscu	
26	207	18.5	1186	92	HSAL32337	Homo sapi	
27	207	18.5	2462	88	AF145440	Homo sapi	
28	207	18.5	2544	88	AF145439	Homo sapi	
29	207	18.5	2577	93	HSUA5982	Homo sapi	
c	30	207	18.5	176968	86	AC005669	Homo sapi
31	206.4	18.4	2751	9	ARI07244	Sequence	
32	201.2	18.0	206782	69	AC024150	Homo sapi	
33	198.2	17.7	1631	94	AF121670	Rattus no	
34	194.6	17.4	1744	7	BTU19947	Bos taurus	
35	184	16.4	1068	88	AF127528	Callithri	
36	182.4	16.3	2214	93	HUMH145	Human mRNA	
37	181.4	16.2	1495	93	HUMCCCKRIA	L09230 Human C-C c	
38	181.4	16.2	1609	10	E13385	cDNA encodi	
39	181.4	16.2	2156	10	I58541	Sequence 1	
40	181.4	16.2	2156	93	HUMRANTES	L10918 Homo sapien	
c	41	181.4	16.2	195028	69	AC024739	Homo sapi
42	178.8	16.0	1473	7	AF047047	Canis fam	
43	177.6	15.9	1068	7	RAB1L8C	M74240 Oryctolagus	
44	177.6	15.9	1232	7	RAB1L8REC	M82873 Oryctolagus	
45	177.4	15.9	1269	85	AB023890	Homo sapi	

1255 bp

DNA

PRI

13-NOV-1996

HSCKRL3

H.sapiens G protein-coupled receptor CKR-L3.

279784

279784.1

GI:1668737

G Protein-coupled Receptor CKR-L3.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 1255)

Gutierrez, J., Varona, R., Zaballos, A., Lind, P. and Marquez, G.

unpublished

unpublished

2 (bases 1 to 1255)

Zaballos, A.

Direct Submission

Submitted (03-SEP-1996) Angel Zaballos, Research, Pharmacia & Upjohn, Antonio Lopez 109, Madrid, 28026, Spain

3 (bases 1 to 1255)

Zaballos, A., Varona, R., Gutierrez, J., Lind, P. and Marquez, G.

Molecular cloning and RNA expression of two new human chemokine receptor-like genes

Biochem. Biophys. Res. Commun. 227 (3), 846-853 (1996)

97040707

Erratum: [[published erratum appears in Biochem Biophys Res Commun 1997 Feb 13;231(2):519-20]]

Location/Qualifiers

1. .1255

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="6"

76. .1185

1

HSCKRL3

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

FEATURES

source

CDS

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1</

BASE COUNT	302 a	294 c	284 g	375 t
ORIGIN				
Query Match	98.4%;	Score 1100.6;	DB 92;	Length 1255;
Best Local Similarity	99.6%;	Pred. No. 2e-307;		
Matches 1114;	Conservative	0;	Mismatches	4;
Indels	1;	Gaps	1;	
QY	1	ATGTTTTCGACTCCAGTGAAGATTATTTTGTGTCAGTCAATACTTATCTATCTACTGAGTTG	60	
Db	89	ATGTTTTCGACTCCAGTGAAGATTATTTTGTGTCAGTCAATACTTATCTATCTACTGAGTTG	148	
QY	61	ATTCTCAGATGTTACTGTGCTCCTTCGACGAGTCCAGGAGTCTCCAGGCTATTTGTAC	120	
Db	149	ATTCTCAGATGTTACTGTGCTCCTTCGACGAGTCCAGGAGTCTCCAGGCTATTTGTAC	208	
QY	121	CGAATTGCCCTACTCCCTTGATCTGTGTTGCTTGGCTCCTCGGGAATATCTGGTGGTGATC	180	
Db	209	CG-ATTGCCCTACTCCCTTGATCTGTGTTGCTTGGCTCCTCGGGAATATCTGGTGGTGATC	267	
QY	181	ACGTTTGGCTTTTATGAAGGCGGAGTCTATGACAGCTCTATCTCTTTGAACATGGCC	240	
Db	268	ACGTTTGGCTTTTATGAAGGCGGAGTCTATGACAGCTCTATCTCTTTGAACATGGCC	327	
QY	241	ATTGCAGACATCTCTTTGTTTCTTACTCTCCCATTCGGGCGAGTGCATGCCAGTGGT	300	
Db	328	ATTGCAGACATCTCTTTGTTTCTTACTCTCCCATTCGGGCGAGTGCATGCCAGTGGT	387	
QY	301	GGTGGGTTTTCAGCAATGCCAGTGCAGTTGCTAAAGGCATCTATGCCATCAACTTT	360	
Db	388	GGTGGGTTTTCAGCAATGCCAGTGCAGTTGCTAAAGGCATCTATGCCATCAACTTT	447	
QY	361	AACTCGGGATGCTGCTCCTGACTTGCATTAGCATGACCGGTACATCGCCATTGTTACAG	420	
Db	448	AACTCGGGATGCTGCTCCTGACTTGCATTAGCATGACCGGTACATCGCCATTGTTACAG	507	
QY	421	GGGACTAAGTCAATCCGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCACTGCGTT	480	
Db	508	GGGACTAAGTCAATCCGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCACTGCGTT	567	
QY	481	GTGTGTGGGGGTGTGTCAGTCAATCCTCCAGTCAACTTTTGTCTTCAACCAAAAATAC	540	
Db	568	GTGTGTGGGGGTGTGTCAGTCAATCCTCCAGTCAACTTTTGTCTTCAACCAAAAATAC	627	
QY	541	AACACCAAGGCGGATGCTGCTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGG	600	
Db	628	AACACCAAGGCGGATGCTGCTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGG	687	
QY	601	TGGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTTATCCCTTTGATGTTCC	660	
Db	688	TGGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTTATCCCTTTGATGTTCC	747	
QY	661	ATGATATTTTGTACAGTTTCATTGTCAAAACCTTGTCGAAGCTCAGATCTCTAAAGG	720	
Db	748	ATGATATTTTGTACAGTTTCATTGTCAAAACCTTGTCGAAGCTCAGATCTCTAAAGG	807	
QY	721	CACAAAGCCATCCGTTGATATCATAGCTGTGCTGCTTGTGTTTCTGGCTTCTGAGATTCCT	780	
Db	808	CACAAAGCCATCCGTTGATATCATAGCTGTGCTGCTTGTGTTTCTGGCTTCTGAGATTCCT	867	
QY	781	CATAACATGGTCTGCTTGTGACGGCTGCTAATTTGGGTAAATGAACCGATCCTGCCAG	840	

Db 868 CATAACATGGCTCTGTGACGGCTGCAATTTGGGTAAATGAACGATCCTGCCAG 927  
QY 841 AGCGAAAAGCTAATGGCTATACGAAAAGCTGTCACAGAAGTCTCGCTTCTGCACTGC 900  
Db 928 AGCGAAAAGCTAATGGCTATACGAAAAGCTGTCACAGAAGTCTCGCTTCTGCACTGC 987  
QY 901 TGCCTGAACCTGTGCTCTACGCTTTTATGGGAGAGTTCAGAAACTACTTTCTGAAG 960  
Db 988 TGCCTGAACCTGTGCTCTACGCTTTTATGGGAGAGTTCAGAAACTACTTTCTGAAG 1047  
QY 961 ATCTTGAAGGACCTGTGCTGTGAGAGGAGTACAAGTCTCAGGCTTCTCCTGTGCC 1020  
Db 1048 ATCTTGAAGGACCTGTGCTGTGAGAGGAGTACAAGTCTCAGGCTTCTCCTGTGCC 1107  
QY 1021 GGGAGGTACTCAGAAAACATTTCTCGGAGACCTGAGACCGCAGATAACCAATCGG 1080  
Db 1108 GGGAGGTACTCAGAAAACATTTCTCGGAGACCTGAGACCGCAGATAACCAATCGG 1167  
QY 1081 TCGTCTTCACTATGTATAGAAAGCTGAGTCTCCCTAA 1119  
Db 1168 TCGTCTTCACTATGTATAGAAAGCTGAGTCTCCCTAA 1206  
RESULT 2  
LOCUS HSU45984 3693 bp DNA PRI 09-JUL-1997  
DEFINITION Homo sapiens CCR6 chemokine receptor (CMKBR6) gene, complete cds.  
ACCESSION U45984  
VERSION U45984.1 GI:2246432  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3693)  
Baba.M., Imai.T., Nishimura.M., Kakizaki.M., Takagi.S.,  
Hieshima.K., Nomiya.H. and Yoshie.O.  
Identification of CCR6, the specific receptor for a novel  
lymphocyte-directed CC chemokine LARC  
J. Biol. Chem. 272 (23), 14893-14898 (1997)  
97313465  
2 (bases 1 to 3693)  
Lautens.L., Modi.W. and Bonner.T.I.  
Cloning, Tissue Distribution and Chromosomal Localization of a  
potential G-Protein-Linked Receptor  
Unpublished  
3 (bases 1 to 3693)  
Bonner.T.I.  
Direct Submission  
Submitted (16-JAN-1996) Tom I. Bonner, Lab of Cell Biology, NIMH,  
Bldg 36, Rm 3A-17, MSC 4090, Bethesda, MD 20892-4090, USA  
On Jul 8, 1997 this sequence version replaced gi:1245058.  
Location/Qualifiers  
1..3693  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="6q27"  
/clone="GPR-CY4"  
join(35..140,237..3136)  
132..1352  
/gene="CMKBR6"  
join(132..140,237..1352)  
/note="G protein-coupled receptor"  
/codon\_start=1  
/product="CCR6 chemokine receptor"  
/protein\_id="AAB62714.1"  
/db\_xref="GI:2251211"  
SRFPVPIAYSLICVGLNLIIVITAFYKARSMTDVLNNAIADILFVLTPFW  
AVSHATGAWFVSATCKLLKGIYAINFCNGLLTCISMDRYIAIVQATSFRLSRT

LPKSKICLVVWGLSVIISSTFVFNQKNTQSDYCEPKYQTVSEPIRWKLLMGLLE  
LTFGFFIPLMFIFCYTIVTKLVQNSKRHKAIRVIAVVLVFLACQIPHNNVLLV  
TAANLGMNRSCQSEKLIQYTKVTEVLAFLHCLLNPVLYAFIQGRFNYFLKILKDL  
WCVRRKYKSSGSCAGRYSENISRQTSADNDNASFTM"  
polyA\_signal 3113..3118  
polyA\_site 3137  
BASE COUNT 1015 a 726 c 828 g 1124 t  
ORIGIN  
Query Match 98.4%; Score 1100.6; DB 93; Length 3693;  
Best Local Similarity 99.6%; Pred. No. 2.3e-40;  
Matches 1114; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 1 ATCTTTTCGACTCCAGTGAAGATTATTTCTGTCAGTCAATCTCTATATTTACTCAGTTG 60  
Db 256 ATCTTTTCGACTCCAGTGAAGATTATTTCTGTCAGTCAATCTCTATATTTACTCAGTTG 315  
QY 61 ATCTCTGAGATGTTACTGTCTCTCTGAGGAGGTTCAGGAGTTCCTCCAGGCTATTTGTAC 120  
Db 316 ATCTCTGAGATGTTACTGTCTCTCTGAGGAGGTTCAGGAGTTCCTCCAGGCTATTTGTAC 375  
QY 121 CGAATTGCTCTCTGATCTCTGTTCTTGGCTCTCTGCGGAAATATTTCTGGTGGTATC 180  
Db 376 CG-ATTGCTCTCTGATCTCTGTTCTTGGCTCTCTGCGGAAATATTTCTGGTGGTATC 434  
QY 181 ACCTTTGCTTTTATAGAGGCGAGTCTATGACAGACGCTCTATCTTTGAAATGCCC 240  
Db 435 ACCTTTGCTTTTATAGAGGCGAGTCTATGACAGACGCTCTATCTTTGAAATGCCC 494  
QY 241 ATTCGACAGATCCTCTTTGTTTACTCTCCCATTTCTGGGAGTGCATGCCACTGGT 300  
Db 495 ATTCGACAGATCCTCTTTGTTTACTCTCCCATTTCTGGGAGTGCATGCCACTGGT 554  
QY 301 GCGTGGGTTTTCAGCAATGCCAGTTCGCAAGTTGCTAAAGGCACTATGCCATCAACTTT 614  
Db 555 GCGTGGGTTTTCAGCAATGCCAGTTCGCAAGTTGCTAAAGGCACTATGCCATCAACTTT 614  
QY 361 AACTGCGGGATGCTGCTCTGACTTGCATTTAGTAGTGCACGGTACATCGCATTTGACAG 420  
Db 615 AACTGCGGGATGCTGCTCTGACTTGCATTTAGTAGTGCACGGTACATCGCATTTGACAG 674  
QY 421 GCGACTAAGTCATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCCTT 480  
Db 675 GCGACTAAGTCATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCCTT 734  
QY 481 GTTGTGTGGGGCTGTGAGTCACTATCTCCAGTCAACTTTTGTCTTCAACCAAAATAC 540  
Db 735 GTTGTGTGGGGCTGTGAGTCACTATCTCCAGTCAACTTTTGTCTTCAACCAAAATAC 794  
QY 541 AACACCCAGGCGAGTCTGTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGG 600  
Db 795 AACACCCAGGCGAGTCTGTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGG 854  
QY 601 TGGAACTGCTGATGTTGGGGCTTTGAGTACTCTTTGTTGTTTCTTATCCCTTTGATGTC 660  
Db 855 TGGAACTGCTGATGTTGGGGCTTTGAGTACTCTTTGTTGTTTCTTATCCCTTTGATGTC 914  
QY 661 ATGATATTTGTTTACAGTTTCATTGTCAAAACCTTTGTCGAAGTCAAGATTTCTAAAAG 720  
Db 915 ATGATATTTGTTTACAGTTTCATTGTCAAAACCTTTGTCGAAGTCAAGATTTCTAAAAG 974  
QY 721 CACAAAGCCATCCGTGTATCATAGTGTGGTGTCTTGTGTTCTTGGCTTGTTCAGATTCCT 780  
Db 975 CACAAAGCCATCCGTGTATCATAGTGTGGTGTCTTGTGTTCTTGGCTTGTTCAGATTCCT 1034  
QY 781 CATAACATGCTCTGCTGTGAGGCTGCTAAATTTGGGTGAAATGAACCGATCTGCCAG 840  
Db 1035 CATAACATGCTCTGCTGTGAGGCTGCTAAATTTGGGTGAAATGAACCGATCTGCCAG 1094  
QY 841 AGCGAAAAGCTAATTTGGCTATAGAAAACCTGTACAGAAAGTCTCTGGCTTTCTCCTGCATGC 900  
Db 1095 AGCGAAAAGCTAATTTGGCTATAGAAAACCTGTACAGAAAGTCTCTGGCTTTCTCCTGCATGC 1154

QY 901 TGCCTGAACCCCTGCTCTACACCTTTTATTGGCAGAAAGTTTCAGAACTACTTTCTGAAG 960  
 |||||  
 Db 1155 TGCCTGAACCCCTGCTCTACACCTTTTATTGGCAGAAAGTTTCAGAACTACTTTCTGAAG 1214  
 |||||  
 QY 961 ATCTTTGAAGGACCTGTGGTGTGTGAGAGGAAGTACAAGTCTCTCAGGCTTCCTCTGTGCC 1020  
 |||||  
 Db 1215 ATCTTTGAAGGACCTGTGGTGTGTGAGAGGAAGTACAAGTCTCTCAGGCTTCCTCTGTGCC 1274  
 |||||  
 QY 1021 GGGAGTACTCAGAAAACATTTCTCGCAGACACGAGTGAAGACCGCAGATTAACGACATGCG 1080  
 |||||  
 Db 1275 GGGAGTACTCAGAAAACATTTCTCGCAGACACGAGTGAAGACCGCAGATTAACGACATGCG 1334  
 |||||  
 QY 1081 TCGTCTTCACTATGTGATAGAAAGCTAGTCTCCCTAA 1119  
 |||||  
 Db 1335 TCGTCTTCACTATGTGATAGAAAGCTAGTCTCCCTAA 1373  
 |||||

RESULT 3  
 HBSA517H2  
 LOCUS HBSA517H2 101230 bp DNA PRI 24-JUL-2000  
 DEFINITION Human DNA sequence from clone RP11-517H2 on chromosome 6 Contains the CCR6 (chemokine (C-C motif) receptor 6) gene, the GPR31 (G protein-coupled receptor 31) gene, the TCP10 (t-complex 10 (a murine tcp homolog)) gene, ESTs, STSS and GSSs, complete sequence.  
 ACCESSION AL121935.17 GI:8247260  
 VERSION AL121935.17  
 KEYWORDS HTG; CCR6; GPR31; receptor; TCP10.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 101230)  
 AUTHORS Almeida, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 COMMENT requests: clonerequests@sanger.ac.uk  
 On Jun 4, 2000 this sequence version replaced gi:8218074.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sv., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>  
 RP11-517H2 is from the library RPCI-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>  
 VECTOR: pBACe3.6  
 IMPORTANT: This sequence is not the entire insert of clone RP11-517H2 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.  
 The true right end of clone RP11-517H2 is at 101230 in this sequence. The true right end of clone RP1-167A14 is at 100 in this sequence.  
 Location/Qualifiers  
 1. .101230  
 /organism="Homo sapiens"

/db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="RP11-517H2"  
 /clone\_lib="RPCI-11.2"  
 1404. .1726  
 /note="LTR16C repeat: matches 56. .387 of consensus"  
 1722. .1916  
 /note="MER39b repeat: matches 13. .224 of consensus"  
 1858. .1969  
 /note="LTR29 repeat: matches 1. .118 of consensus"  
 2602. .2956  
 /note="MLT1A1 repeat: matches 1. .361 of consensus"  
 2957. .3263  
 /note="MLT1B repeat: matches 128. .451 of consensus"  
 3261. .3310  
 /note="MLT1B repeat: matches 128. .175 of consensus"  
 3308. .3472  
 /note="MLT1B repeat: matches 1. .168 of consensus"  
 5304. .5844  
 /note="match: GSS: Em:AQ549970"  
 5429. .5511  
 /note="AluJb repeat: matches 53. .135 of consensus"  
 6614. .6693  
 /note="MADE1 repeat: matches 1. .80 of consensus"  
 7055. .8158  
 /note="L2 repeat: matches 511. .1507 of consensus"  
 8159. .8474  
 /note="AluSg repeat: matches 1. .310 of consensus"  
 8480. .9050  
 /note="L1MA5A repeat: matches 5711. .6294 of consensus"  
 9051. .9917  
 /note="L2 repeat: matches 1510. .2745 of consensus"  
 9940. .10176  
 /note="MIR repeat: matches 20. .261 of consensus"  
 15165. .15852  
 /note="match: GSS: Em:AQ374466"  
 15905. .16006  
 /note="MER81 repeat: matches 1. .113 of consensus"  
 16268. .16357  
 /note="3 copies 30 mer 81% conserved"  
 16269. .16340  
 /note="3 copies 24 mer 80% conserved"  
 16506. .16815  
 /note="AluI repeat: matches 1. .311 of consensus"  
 17039. .17066  
 /note="14 copies 2 mer tt 89% conserved"  
 17763. .17944  
 /note="MIR repeat: matches 20. .202 of consensus"  
 19526. .19588  
 /note="3 copies 21 mer 82% conserved"  
 20849. .21230  
 /note="MER92A repeat: matches 34. .412 of consensus"  
 21618. .21856  
 /note="MIR repeat: matches 17. .262 of consensus"  
 22931. .23186  
 /note="AluSx repeat: matches 1. .256 of consensus"  
 23316. .23437  
 /note="61 copies 2 mer aa 63% conserved"  
 23777. .23884  
 /note="MIR repeat: matches 131. .245 of consensus"  
 23892. .23981  
 /note="3 copies 30 mer 77% conserved"  
 23894. .23965  
 /note="3 copies 24 mer 81% conserved"  
 23896. .23979  
 /note="3 copies 28 mer 78% conserved"  
 23901. .23976  
 /note="38 copies 2 mer ca 72% conserved"  
 23982. .24233  
 /note="9 copies 28 mer 61% conserved"  
 23985. .24128  
 /note="6 copies 24 mer 71% conserved"  
 24038. .24127

FEATURES  
 source



QY 1021 GGGAGGTACTCAGAAAACATTTCTCGGAGACCCAGTGGAGACCGCAGATACGACAAATGGG 1080  
 |||||||  
 Db 36379 GGGAGGTACTCAGAAAACATTTCTCGGAGACCCAGTGGAGACCGCAGATACGACAAATGGG 36438  
 |||||||  
 QY 1081 TCGTCTTCACTATGTGATAGAAAGCTGAGTCTCCCTAA 1119  
 |||||||  
 Db 36439 TCGTCTTCACTATGTGATAGAAAGCTGAGTCTCCCTAA 36477  
 |||||||

RESULT 4  
 HSD68032  
 LOCUS HSD68032 1518 bp DNA PRI 06-MAR-1997  
 DEFINITION Human G protein-coupled receptor (STRL22) gene, complete cds.  
 ACCESSION U68032  
 VERSION U68032.1 GI:1870668  
 KEYWORDS  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1518)  
 AUTHORS Liao, F., Lee, H. H. and Farber, J. M.  
 TITLE Cloning of STRL22, a new human gene encoding a G-protein-coupled receptor related to chemokine receptors and located on chromosome 6q27

JOURNAL Genomics 40 (1), 175-180 (1997)  
 MEDLINE 97224503

REFERENCE 2 (bases 1 to 1518)  
 AUTHORS Farber, J. M. and Liao, F.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-AUG-1996) NIAID/BDlg10/Rmln-228, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA

FEATURES  
 Location/Qualifiers  
 1..1518  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Stratagene # 951202"  
 /chromosome="6"  
 /map="6q27"  
 /sex="male"  
 /note="Caucasian"  
 1..1518  
 /gene="STRL22"  
 <1..2  
 /gene="STRL22"  
 3..108  
 /gene="STRL22"  
 join(100..108,205..1320)  
 /gene="STRL22"  
 /note="contains seven transmembrane domains; expressed in lymphocytes"  
 /codon\_start=1  
 /product="G protein-coupled receptor"  
 /protein\_id="AAC51125.1"  
 /db\_xref="GI:1870669"  
 /translation="MSGESMNFSDVDSSEDFVSVNTSYVSDSEMLCSLOEVROF  
 SRLFPVIALICVFLGLNLVITAFYKARSMTDVYLLNMAIADLLFVLTLPFW  
 AVSHATGAWFSNATCKLLKGIYALNFNGMLLLTCISMDRYIALVOATKSFRLSPFT  
 LPRKILCLVWVGLVLISSSTVFVFNQKYNKTOGSDVCEPKYQTVSEPIRWKLLMLGLE  
 LLFGFFPLMFMIICYTFIVKTLVQAKSRKRAIRVIAVVLVFLACQIPHNWVLLV  
 TAAHLGKMRSCQSEKLLIGYTKTVTEVLAFLHCCILNPVLYAFIGQKFRNYFLKILKDL  
 WCVRKRYKSGFSFCAGRYSENISROTSETADNDNASSFTM"  
 109..204  
 /gene="STRL22"  
 205..>1518  
 /gene="STRL22"  
 360 a 354 c 351 g 453 t

BASE COUNT 360 a 354 c 351 g 453 t  
 ORIGIN

Query Match 98.1%; Score 1097.4; DB 93; Length 1518;  
 Best Local Similarity 99.4%; Pred. No. 1.7e-306;

Matches 1112; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
 QY 1 ATGTTTTGAGTCCAGTGAAGATTATTTGTGTCACTCAATACTTATATTTACTCAGTTG 60  
 |||||||  
 Db 224 ATGTTTTGAGTCCAGTGAAGATTATTTGTGTCACTCAATACTTATATTTACTCAGTTG 283  
 |||||||  
 QY 61 ATTCTGAGATGTTACTGTGCTCCTTGCAGGAGGTCCAGCAGTTCTCCAGGCTATTGTGAC 120  
 |||||||  
 Db 284 ATTCTGAGATGTTACTGTGCTCCTTGCAGGAGGTCCAGCAGTTCTCCAGGCTATTGTGAC 343  
 |||||||  
 QY 121 CGAATTCCTACTCCTTGTGATCTGTGCTTGGCCCTCCCTGGGGAATATCTCGTGGTGATC 180  
 |||||||  
 Db 344 CG-ATTGCCCTACTCCTTGTGATCTGTGCTTGGCCCTCCCTGGGGAATATCTCGTGGTGATC 402  
 |||||||  
 QY 181 ACCTTTGCTTTTTATAAGAGGCCAGGTCTATGACAGACGTCTATCTCTTGAACATGGCC 240  
 |||||||  
 Db 403 ACCTTTGCTTTTTATAAGAGGCCAGGTCTATGACAGACGTCTATCTCTTGAACATGGCC 462  
 |||||||  
 QY 241 ATTCCAGACATCCCTTTGTTTACTCTCCATTTCTGGGAGTGGAGTCAATGCCACTGGT 300  
 |||||||  
 Db 463 ATTCCAGACATCCCTTTGTTTACTCTCCATTTCTGGGAGTGGAGTCAATGCCACTGGT 522  
 |||||||  
 QY 301 GCGTGGGTTTTACCAATGCCAGTGCAGTGTGCTTAAAGGCATCTATGCCATCAACTTT 360  
 |||||||  
 Db 523 GCGTGGGTTTTACCAATGCCAGTGCAGTGTGCTTAAAGGCATCTATGCCATCAACTTT 582  
 |||||||  
 QY 361 AACTGCGGGATGCTGCTCCTGACTTGCATTAGCATGGACCGGTACATGCCATTTGTACAG 420  
 |||||||  
 Db 583 AACTGCGGGATGCTGCTCCTGACTTGCATTAGCATGGACCGGTACATGCCATTTGTACAG 542  
 |||||||  
 QY 421 GCGACTAAGTCAATCCGGCTCCGATCCAGAACACTTACCGCGAGCAAAATCATCTGCCTT 480  
 |||||||  
 Db 643 GCGACTAAGTCAATCCGGCTCCGATCCAGAACACTTACCGCGAGCAAAATCATCTGCCTT 702  
 |||||||  
 QY 481 GTTGTGTTGGGGCTGCTCAGTCAATCTCCAGTCAACTTTTGTCTTCAACCAAAATAC 540  
 |||||||  
 Db 703 GTTGTGTTGGGGCTGCTCAGTCAATCTCCAGTCAACTTTTGTCTTCAACCAAAATAC 762  
 |||||||  
 QY 541 AACACCCAAAGGCAGCGATGCTGTGAACCCCAAGTACCAAACTGTCTCGAGGCCATCAGG 600  
 |||||||  
 Db 763 AACACCCAAAGGCAGCGATGCTGTGAACCCCAAGTACCAAACTGTCTCGAGGCCATCAGG 822  
 |||||||  
 QY 601 TGAAGCTGCTGATGTTGGGGCTTGAGTACTCTTTTGGTTTCTTTATCCCTTTGATGTTTC 660  
 |||||||  
 Db 823 TGAAGCTGCTGATGTTGGGGCTTGAGTACTCTTTTGGTTTCTTTATCCCTTTGATGTTTC 882  
 |||||||  
 QY 661 ATGATATTTTGTACAGTTCATTTGTCAAAACCTTGGTGCAGCTCAGAAATCTAAAGG 720  
 |||||||  
 Db 883 ATGATATTTTGTACAGTTCATTTGTCAAAACCTTGGTGCAGCTCAGAAATCTAAAGG 942  
 |||||||  
 QY 721 CACAAAGCCATCCCTGTGTAATCATAGCTGTGGTGTGTTTGTGTTTCTGGCTTGTGAGATTCT 780  
 |||||||  
 Db 943 CACAAAGCCATCCCTGTGTAATCATAGCTGTGGTGTGTTTGTGTTTCTGGCTTGTGAGATTCT 1002  
 |||||||  
 QY 781 CATAACATGTTCTGCTTGTGAGCGGCTGCTAATTTGGTGTAAATGAACCGATCTCGCCAG 840  
 |||||||  
 Db 1003 CATAACATGTTCTGCTTGTGAGCGGCTGCTAATTTGGTGTAAATGAACCGATCTCGCCAG 1062  
 |||||||  
 QY 841 AGCCAAAAGCTAATTTGGCTATACCAAACTCTCACAGAAAGTCTCGGCTTTCTCGCAGTGC 900  
 |||||||  
 Db 1063 AGCCAAAAGCTAATTTGGCTATACCAAACTCTCACAGAAAGTCTCGGCTTTCTCGCAGTGC 1122  
 |||||||  
 QY 901 TGCCTGAACCTGTGCTCTACGCTTTTATTGGGAGAGTTCAGAAAATCTACTTCTCGAAG 960  
 |||||||  
 Db 1123 TGCCTGAACCTGTGCTCTACGCTTTTATTGGGAGAGTTCAGAAAATCTACTTCTCGAAG 1182  
 |||||||  
 QY 961 ATCTTGAAGGACCTGTGGTGTGAGAGGAGTACAAAGTCTCCTCAGGCTTCTCTGTGTC 1020  
 |||||||  
 Db 1183 ATCTTGAAGGACCTGTGGTGTGAGAGGAGTACAAAGTCTCCTCAGGCTTCTCTGTGTC 1242  
 |||||||  
 QY 1021 GGGAGGTACTCAGAAAACATTTCTCGGAGACCCAGTGGAGACCGCAGATACGACAAATGGG 1080  
 |||||||  
 Db 1243 GGGAGGTACTCAGAAAACATTTCTCGGAGACCCAGTGGAGACCGCAGATACGACAAATGGG 1302  
 |||||||

```

QY 1081 TCGCTCTTCACTATGTGATAGAAAGCTGAGTCTCCCTAA 1119
|||||
Db 1303 TCGCTCTTCACTATGTGATAGAAAGCTGAGTCTCCCTAA 1341
|||||

RESULT 5
HSU68030 2978 bp mRNA PRI 06-MAR-1997
LOCUS Human G protein-coupled receptor (STRL22) mRNA, complete cds.
DEFINITION
ACCESSION U68030
VERSION U68030.1 GI:1870665
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2978)
Liao,F., Lee,H.H. and Farber,J.M.
AUTHORS Cloning of STRL22, a new human gene encoding a G-protein-coupled
TITLE receptor related to chemokine receptors and located on chromosome
6q27
JOURNAL Genomics 40 (1), 175-180 (1997)
MEDLINE 9724503
REFERENCE 2 (bases 1 to 2978)
AUTHORS Farber,J.M. and Liao,F.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-1996) NIAID, Bldg. 10, Rm. 11N-228, NIH, 9000
Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
    location/Qualifiers
    1..2978
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="6"
        /map="6q27"
        /cell_type="tumor infiltrating lymphocyte"
    1..264
        /gene="STRL22"
        /note="one of two alternatively spliced forms that differ
        in the 5'UTR; see GenBank Accession Number U68031"
    1..1389
        /gene="STRL22"
        /note="contains seven transmembrane domain receptors"
    /product="G protein-coupled receptor"
    /protein_id="AAC51124.1"
    /db_xref="GI:1870666"
    /translation="MSGESMNFSDVFDSDSEDFVSVNTSYSDSEMLLCSLOEVROF
    SRLFVPIAYSLICVGLLGNLIVIPAFYKARSMTDVLNMAIADILFVLTPFW
    AVSHAYGAWVFSNATCKLLKGIYAINFNCGMLLITCISMDRYIAIVQATKSFRLRSRT
    LPRTICLIVWVLSVLISSSTFVNOKYNTGSDVCEPKYQIVSEPIRWKLLMLGLE
    LFGFFPIPLMFICYFTFVKTIVQAKNSRKAIRVIAVLVFLACQIPHNMLLV
    TANLGMNRSQSEKLIQYTKTVEVLAFLHCLNPLVAFIQKPRNFVFLKILKDL
    WCVRVRYKSGFSFGCSRYSENISROTSETADNDNASFTM"
BASE COUNT 807 a 608 c 714 g 849 t
ORIGIN

Query Match 98.1%; Score 1097.4; DB 93; Length 2978;
Best Local Similarity 99.4%; Pred. No. 1.9e-306;
Matches 1112; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 ATGTTTTCGACTCCAGTGAAGATTATTTGTGTGAGTCAATACCTTATTTACTCAGTTG 60
|||||
Db 293 ATGTTTTCGACTCCAGTGAAGATTATTTGTGTGAGTCAATACCTTATTTACTCAGTTG 352
|||||

QY 61 ATTCGAGATGTTACTGTGCTCTTCGAGGAGGTCAGGAGCTTCTCCAGGCTATTTGTAC 120
|||||
Db 353 ATTCGAGATGTTACTGTGCTCTTCGAGGAGGTCAGGAGCTTCTCCAGGCTATTTGTAC 412
|||||
QY 121 CGAATTCCTACTCTTGAATCTGTGTCTTGGCCCTCTGGGGAATATTTCTGGTGTGATC 180
|||||

```

RESULT 6  
 HSU60000  
 LOCUS HSU60000 1137 bp mRNA PRI  
 DEFINITION Human IL8-related receptor (DRV6) mRNA, complete cds.  
 ACCESSION U60000

VERSION U60000.1 GI:1515434  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1137)  
AUTHORS McCoy.R. and Perlmuter,D.H.  
TITLE Cloning of novel IL8-related receptors from hepatic tissue  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1137)  
AUTHORS McCoy.R. and Perlmuter,D.H.  
TITLE Direct Submission  
JOURNAL Submitted (28-MAY-1996) Pediatrics, Washington University in St.  
Louis, 4942 Parkview Place, St Louis, MO 63110, USA  
FEATURES  
source  
1. 1137  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="hepatoma HepG2"  
1. 1137  
/gene="DRY6"  
1. 1137  
/gene="DRY6"  
/codon\_start=1  
/product="IL8-related receptor"  
/protein\_id="AAB06949.1"  
/db\_xref="GI:1515435"  
/translation="MSGESMNFSDVDSSEDFYVSNTSYYSVDSEMLLCSLOEVRQF  
SRLEVPYSLICVFALLGNLLVITFAFNKKARSMIDVILVNNAIADILFLVLTLPFW  
AVSHATGAWFNSATCLKLLGIYAINENCGMLLTICISMDRYIAIVQATKSFRLRSPT  
LPSKIIICLVWGLUSVLISSSFVFNOKYINFLGSDVCEPKYQTVSPVIRWLLMLGLE  
LLGFFPLPMFICFYIIVKTLVQAGNSKRHKARIVIAVFLVFLACQIPHNMLLV  
TAANLGRNRSQSEKILGYTKTVTEVLAFLHCLNPLVYAFIQGKFRNYFLKILKDL  
WCVRKYKSGFCAGRYSENISROTSETADNDNAVHLHVYIES"  
BASE COUNT 263 a 270 c 266 g 338 t  
ORIGIN  
Query Match 95.2%; Score 1065.2; DB 93; Length 1137;  
Best Local Similarity 98.6%; Pred. No. 3.5e-297;  
Matches 1095; Conservative 0; Mismatches 13; Indels 2; Gaps 2;  
Qy 1 ATGTTTTCAGTCCAGTGAAGATATTTTGTGTCAGTCAATCTCATATCTACTAGTTG 60  
Db 29 ATGTTTTCAGTCCAGTGAAGATATTTTGTGTCAGTCAATCTCATATCTACTAGTTG 88  
Qy 61 ATTCAGATGTTACTGTCTCTTGCAGGAGTTCAGGAGTTCCTCCAGGCTATTGTGAC 120  
Db 89 ATTCAGATGTTACTGTCTCTTGCAGGAGTTCAGGAGTTCCTCCAGGCTATTGTGAC 148  
Qy 121 CGAATTGCTACTCTCTGATCTGCTTCTTGCCTCCCTGGGGAATATTCTGTTGTTGATC 180  
Db 149 CG-ATTGCTACTCTCTGATCTGCTTCTTGCCTCCCTGGGGAATATTCTGTTGTTGATC 207  
Qy 181 ACCTTTGTCTTTATAGAAGCCAGGCTATGACAGACGCTATCTCTTGAACATGGCC 240  
Db 208 ACCTTTGTCTTTATAGAAGCCAGGCTATGACAGACGCTATCTCTGTTGAACATGGCC 267  
Qy 241 ATTCAGACATCCCTTTTGTCTTACTCTCCATTTCTGGGAGTGCATGCCACTGGT 300  
Db 268 ATTCAGACATCCCTTTTGTCTTACTCTCCATTTCTGGGAGTGCATGCCACTGGT 327  
Qy 301 CGGTGGGTTTTTCAGCAATGCCAGTGCAGTTCGCTTAAAGGCAATCTATGCCATCAACTTT 360  
Db 328 CGGTGGGTTTTTCAGCAATGCCAGTGCAGTTCGCTTAAAGGCAATCTATGCCATCAACTTT 387  
Qy 361 AACTCGGGGATGCTGCTCCTGACTTGCATTAGCATGGACCGGTACATGCCATTTGTACAG 420  
Db 388 AACTCGGGGATGCTGCTCCTGACTTGCATTAGCATGGACCGGTACATGCCATTTGTACAG 447  
Qy 421 CGGACTAAGTCAATCCGGCTCGGATCCAGAACACTACCGCGCAGCAAAATCATCTGCCTT 480  
Db 448 GCGACTAAGTCAATCCGGCTCGGATCCAGAACACTACCGCGCAGCAAAATCATCTGCCTT 507  
Qy 481 GTTGTGTGGGGGTGTGTCAGTCAATCTCCAGTCAAACTTTTGTCTTCAACCAAAAATAC 540  
Db 508 GTTGTGTGGGGGTGTGTCAGTCAATCTCCAGTCAAACTTTTGTCTTCAACCAAAAATAC 567  
Qy 541 ACACCCAGGAGGAGGATGCTGTGAACCCCAAGTACCAAACTGTCTCGGAGCCCATCAGG 600  
Db 568 AACACCCCTTGGCAGCGATGCTGTGAACCCCAAGTACCAAACTGTCTCGGAGCCCATCAGG 627  
Qy 601 TGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTTATCCCTTTGATGTTT 660  
Db 628 TGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTTTCCCTTTGATGTTT 687  
Qy 661 ATGATATTTTGTACAGTTTCATTGTCAAAACCTTGGTGAAGCTCAGAATTTCTAAAGAG 720  
Db 688 ATGATATTTTGTACAGTTTCATTGTCAAAACCTTGGTGAAGCTCAGAATTTCTAAAGAG 747  
Qy 721 CACAAAGCCATCCGTGTAATCATAGCTGTGGTCTTGTGTTTCTGGCTTGTCAAGTTCTT 780  
Db 748 CACAAAGCCATCCGTGTAATCATAGCTGTGGTCTTGTGTTTCTGGCTTGTCAAGTTCTT 807  
Qy 781 CATAACATGGTCTGCTTGTGAGCGCTGCTAAATTTGGTAAATGAACCGATCTGCCAG 840  
Db 808 CATAACATGGTCTGCTTGTGAGCGCTGCTAAATTTGGTAAATGAACCGATCTGCCAG 867  
Qy 841 AGCGAAAGCTAATTTGGCTATACGAAACTGTACAGAAAGTCTGGCTTCTCTGCACCTGC 900  
Db 868 AGCGAAAGCTAATTTGGCTATACGAAACTGTACAGAAAGTCTGGCTTCTCTGCACCTGC 927  
Qy 901 TGCTGAACCTGCTGCTCTACGCTTTTATGGGAGAGTTTCAGAACTACTTTCTGTAAG 960  
Db 928 TGCTGAACCTGCTGCTCTACGCTTTTATGGGAGAGTTTCAGAACTACTTTCTGTAAG 987  
Qy 961 ATCTTGAAGACCTGTGGTGTGTGAGAGGAGTACAGTCTCTCAGGCTTCTCTGTGCTCC 1020  
Db 988 ATCTTGAAGACCTGTGGTGTGTGAGAGGAGTACAGTCTCTCAGGCTTCTCTGTGCTCC 1047  
Qy 1021 GGAGTACTCAGAAACATTTCTGGCAGACCTGAGACCCGAGATACGACAATG-C 1079  
Db 1048 GGAGTACTCAGAAACATTTCTGGCAGACCTGAGACCCGAGATACGACAATG-C 1107  
Qy 1080 GTCGCTCTTCACTATGTAGTAAAGCTGA 1109  
Db 1108 GTCGCTCTTCACTATGTAGTAAAGCTGA 1137  
RESULT 7  
AL353591/c  
LOCUS  
DEFINITION Homo sapiens chromosome 6 clone RP11-366H19, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 18 unordered pieces.  
ACCESSION AL353591  
VERSION AL353591.5 GI:9930866  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 205272)  
AUTHORS Sims,S.  
TITLE Direct Submission  
JOURNAL Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
COMMENT On Aug 27, 2000 this sequence version replaced gi:9863646.  
----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: BA366H19

```

----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 192563 bases at least Q40
Consensus quality: 197287 bases at least Q30
Consensus quality: 200442 bases at least Q20
Insert size: 203572; sum-of-contigs
Insert size: 183107; 13.0% error; agarose-fp
Quality coverage: 3.30x in Q20 bases; sum-of-contigs Quality
coverage: 3.75x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 4259: contig of 4259 bp in length
* 4260 4359: gap of 100 bp
* 4360 26102: contig of 21743 bp in length
* 26103 26202: gap of 100 bp
* 26203 28547: contig of 2345 bp in length
* 28548 28647: gap of 100 bp
* 28648 49653: contig of 21006 bp in length
* 49654 49753: gap of 100 bp
* 49754 80134: contig of 30381 bp in length
* 80135 80234: gap of 100 bp
* 80235 83646: contig of 3412 bp in length
* 83647 83746: gap of 100 bp
* 83747 96061: contig of 12315 bp in length
* 96062 96161: gap of 100 bp
* 96162 103311: contig of 7150 bp in length
* 103312 103411: gap of 100 bp
* 103412 122508: contig of 19097 bp in length
* 122509 122608: gap of 100 bp
* 122609 124966: contig of 2358 bp in length
* 124967 125066: gap of 100 bp
* 125067 142955: contig of 17889 bp in length
* 142956 143055: gap of 100 bp
* 143056 161736: contig of 18681 bp in length
* 161737 161836: gap of 100 bp
* 161837 169189: contig of 7353 bp in length
* 169190 169289: gap of 100 bp
* 169290 175251: contig of 5962 bp in length
* 175252 175351: gap of 100 bp
* 175352 181328: contig of 5977 bp in length
* 181329 181428: gap of 100 bp
* 181429 183863: contig of 2435 bp in length
* 183864 183963: gap of 100 bp
* 183964 195169: contig of 11206 bp in length
* 195170 195269: gap of 100 bp
* 195270 205272: contig of 10003 bp in length.
Location/Qualifiers
1. .205272
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone_lib="RP11-366H19"
/clone_lib="RP11-11.2"
1. .4259
/feature="assembly_fragment:00710"
fragment_chain:1"
4360. .26102
/feature="assembly_fragment:00448"
fragment_chain:1"
26203. .28547
/feature="assembly_fragment:01871"
fragment_chain:1"
28648. .49653
/feature="assembly_fragment:01605"

```

```

misc_feature      fragment_chain:1"
49754. .80134
/feature="assembly_fragment:01641"
fragment_chain:1"
80235. .83646
/feature="assembly_fragment:01500"
fragment_chain:1"
83747. .96061
/feature="assembly_fragment:01099"
fragment_chain:1"
96162. .103311
/feature="assembly_fragment:00756"
fragment_chain:2"
103412. .122508
/feature="assembly_fragment:01078"
fragment_chain:2"
122609. .124966
/feature="assembly_fragment:01675"
fragment_chain:2"
125067. .142955
/feature="assembly_fragment:01165"
fragment_chain:3"
143056. .161736
/feature="assembly_fragment:01021"
fragment_chain:3"
161837. .169189
/feature="assembly_fragment:01200"
fragment_chain:3"
169290. .175251
/feature="assembly_fragment:00119"
175352. .181328
/feature="assembly_fragment:00401"
181429. .183863
/feature="assembly_fragment:01812"
183964. .195169
/feature="assembly_fragment:02127"
195270. .205272
/feature="assembly_fragment:01692"
clone_end:17
vector_side:right"
BASE COUNT      53785 a 47341 c 47491 g 54940 t 1715 others
ORIGIN

Query Match      78.3%; Score 876.6; DB 79; Length 205272;
Best Local Similarity 85.5%; Pred. No. 2.6e-242;
Matches 957; Conservative 0; Mismatches 104; Indels 58; Gaps 2;

QY 1 ATGTTTCGACTCCAGTGAAGATTATTTGTCTCAGTCAATACTTCAATATCTCAGTTG 60
|||||
Db 123291 ATGTTTCGACTCCAGTGAAGATTATTTGTCTCAGTCAATACTTCAATATCTCAGTTG 123232
|||||

QY 61 ATTCGAGATGTTACTGTGCTCTTGCAGGAGTTCAGGAGTTCCTCCAGGCTATTTGTATC 120
|||||
Db 123231 ATTCGAGATGTTACTGTGCTCTTGCAGGAGTTCAGGAGTTCCTCCAGGCTATTTGTATC 123172
|||||

QY 121 CGAATTGCCTACTCCTTGCATCTGTGTTCTTGGCCCTCTCCGGGAATATTCGTGGTGGATC 180
|||||
Db 123171 CG-ATTGCTACTCCTTGCATCTGTGTTCTTGGCCCTCTCCGGGAATATTCGTGGTGGATC 123113
|||||

QY 181 ACCTTTGCTTTTATAAGAGCCAGGCTATACAGAGAGCTCTATCTCTTGACATGGCC 240
|||||
Db 123112 ACCTTTGCTTTTATAAGAGCCAGGCTATACAGAGAGCTCTATCTCTTGACATGGCC 123053
|||||

QY 241 ATTGCAGACATCCTCTTTGTTCTTACTCTCCCATCTCTGGCAGTGTGATGCACTGGT 300
|||||
Db 123052 ATTGCAGACATCCTCTTTGTTCTTACTCTCCCATCTCTGGCAGTGTGATGCACTGGT 122993
|||||

QY 301 CGGTGGGTTTTTCAGCAATGCCACGTGCAAGTTGCTAAAAAGGCATCTATGCCATCAACTTT 360
|||||
Db 122992 CGGTGGGTTTTTCAGCAATGCCACGTGCAAGTTGCTAAAAAGGCATCTATGCCATCAACTTT 122933
|||||

QY 361 AACTGCGGGAGTCTGCTCTCTGACTTGCATTAGCATGGAGCCGGTACATCGCCATTGTACAG 420

```

TITLE Mouse G protein-coupled receptor KY411  
 JOURNAL Published Only in Database (1997) in press  
 FEATURES Location/Qualifiers  
 source 1..1302  
     /organism="Mus musculus"  
     /db\_xref="taxon:10090"  
     147..1250  
 CDS /codon\_start=1  
     /product="G protein-coupled receptor KY411"  
     /protein\_id="B2623776.1"  
     /db\_xref="GI:2668410"  
     /translation="MNSTSEYFGTDDYNTVEYSIPDHGPCSLEEVNRFTKVFVDIA  
     YSLICFGLGNITMYNTAFYAKRSMTDVIYLNNAITDILFLVLPFPWAVTSHKVIC  
     WVPSDALCKLQNTMYNTAFYAKRSMTDVIYLNNAITDILFLVLPFPWAVTSHKVIC  
     VAWVFISIISSPTFFINFKYELQDRDCEPRYSVSEPTIKLLGMLGLEFLFGFTFP  
     LLEPMFCYFLIKTLVQAQNSKHRAIRVVIAVLVFLACQIPHNMLVLLVAVNTGKV  
     GRSCSTKEVLAYTRNVAEYLAFLHCLNLPVLFAFGQKFRNFMKIMKDVCMCRRRKNK  
     MPGFLGKLVSESYISQRTSETVENDNASFTM"  
 314 a 331 c 320 g 337 t  
 BASE COUNT  
 ORIGIN

Query Match	56.4%	Score 631	DB 94	Length 1302
Best Local Similarity	75.4%	Pred. No. 1.6e-171		
Matches 811	Conservative	0	Mismatches 260	Indels 4
Gaps	2			
QY	29	TGTCGTCAGTCAATACTTCATATTA	CTACTCAGTTGATCTCGAGATGTTACTGTGCTCCTTGCA	88
DB	179	TGATTATGACAAACACAGAGTATTATTC	TTCCTCCAGACCATGGGCCATGCTCCCTTAGA	238
QY	89	GGAGGTGAGGCGAGTCTCCAGGCTATT	TGTACCGAAATGGCTACTCCTTGTATCTGTGTTTC	148
DB	239	AGAGGTGAGAAACTTCACCAAGGTATT	TGTGCC - AATTGCTACTCCTTAATATGTGTCT	297
QY	149	TTGCCCTCCTGGGGAATATCTG	TGCTGCTGATCACCTTTGCTTTTATAAGAGCCAGGT	208
DB	298	TTGCCCTCCTGGGCAACATTA	TGTGCTGATGACCTTTGCTCTCTACAAGAAAGACAGAT	357
QY	209	CTATGACAGACGCTCTATCTCT	TGAACATGGCCATTGCAGACATCCCTCTTTGTCTTACTC	268
DB	358	CCATGACTGAGTCTACCTGTT	GAACATGGCCATCAGACATACATCTTTGTCTCACCC	417
QY	269	TCCCATTTCTGGGCGAGTGAGTCA	TGCCACTGGTGGTGGGTTTTTCAGCAATGCCACGTGCA	328
DB	418	TACCGTTCTGGCGAGTTACTCAT	TGCCACCACAACTTGGGTTTTTCAGCGATGCATGTGTA	477
QY	329	AGTTGCTTAAAGGACTATATGCC	ATCAACTTTAACTGCGGGATGCTGCTCCTGACITGCA	388
DB	478	AACTGATGAAGGCACATATAT	TGCGGTCAACTTTAACTGTGGGATGCTGCTCCTGGCCCTGA	537
QY	389	TTAGCATGGACCGGTACATCGCC	ATTCTACAGGCGACTAAAGTCATTTCCCGGCTCCGATFCCA	448
DB	538	TCAGCATGGACCGGTACATTTGCC	ATCTGCCAGCAACCAATCTTTCCGGGTAGGTTCCA	597
QY	449	GAACATACCGCGCGCAAAATCAT	CTGCCTTGTTGTGGGGGCTGTCTAGTCATCATCT	508
DB	598	GAACACTGACACAGTAAGTCTAT	CTGTGGCGAGTGTGTTTCACTCCATCATCATCT	657
QY	509	CCAGCTCAACTTTTGTCTTCA	CCAAAAATACACACCCCAAGCAGGGATGCTGTGTAAC	568
DB	658	CAAGCCCTACATTTATCTTCA	CAAGAANAATACGAGCTGCAAGATCTGTGTCTGTGAGC	717
QY	569	CCAAGTACCAAACTGCTCGG	AGCCCATCAGGTGGAAGCTGCTGATGTTGGGGCTTGAGC	628
DB	718	CACGGTACAGGTCTGTCTC	AGACCCCATCAGTGGGAAGCTGCTGGGTATGGGACTGGAGC	777
QY	629	TACTCTTTGGTCTTTTATCC	CTTTTGATGTTTCATGATATTTTGTATACAGTTCATTGTCA	688
DB	778	TGTTCTTTGGTCTTCTAC	CCCTTTTGTCTTTATGTTGTTCTGCTATCTGTTCAATATCA	837
QY	689	AAACCTTTGGTGAAGCTCAG	AATTTCTTAAAGGCACAAAGCCATCCCGTGAATCATAGCTG	748
DB	838	AGACCTTTGGTGCAGGCC	CCCAACTTCCAAAGAGCACAAGACCATCCCGAGTGTGATCGCTG	897

RESULT	8	
AB009369		
LOCUS	AB009369	1302 bp mRNA
DEFINITION	Mus musculus mRNA for G protein-coupled receptor KY411, complete cds.	09-DEC-1997
ACCESSION	AB009369	ROD
VERSION	AB009369.1	
KEYWORDS	G protein-coupled receptor KY411.	
SOURCE	Mus musculus cDNA to mRNA.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 1302)	
TITLE	Yanagihara, S., Komura, E. and Yamaguchi, Y.	
JOURNAL	Submitted Direct Submission	
	Submitted (01-DEC-1997) to the DDBJ/EMBL/GenBank databases.	
	Shigeniro Yanagihara, Kirin Brewery Co., Ltd., Pharmaceutical Research Laboratory; 3-Miyahara-cho, Takasaki, Gunma 370-12, Japan (E-mail: syanagihara@kirin.co.jp, Tel: 81-273-46-9826)	
REFERENCE	2 (bases 1 to 1302)	
AUTHORS	Yanagihara, S., Komura, E. and Yamaguchi, Y.	





JOURNAL	transactivated by Epstein-Barr virus nuclear antigen 2
MEDLINE	Biochem. Biophys. Res. Commun. 215 (2), 737-743 (1995)
REFERENCE	96011839
AUTHORS	2 (bases 1 to 1500)
TITLE	Lipp, M.
JOURNAL	Direct Submission
FEATURES	Submitted (10-FEB-1995) M. Lipp, Max-Delbrueck-Centrum fuer Mol.Medizin, Robert-Roessle-Strasse 10, 13122 Berlin-Buch, FRG
source	Location/Qualifiers 1..1500

```

intron      <1..403
            /number=1
repeat_region 1..310
            /rpt_family="Alu"
misc_feature 375..405
            /note="splice acceptor site"
exon        404..>1500
            /number=2
gene        404..1480
            /gene="BLR2"
mRNA        <404..1480
            /gene="BLR2"
            /product="Bukitt's lymphoma receptor 2"
BASE COUNT 324 a 457 c 363 g 356 t
ORIGIN

```

Query Match	19.8%;	Score 222;	DB 92;	Length 1500;
Best Local Similarity	55.5%;			
Matches 515:	Conservative	Pred. No. 4e-53;		
	0;	Mismatches 400;	Indels 13;	Gaps 4;

QY	71	GT	TACTGTGCTCCTTGCAGGAGGTCAGGAGTTCTCCAGGCTATTTGTACCGAAATGGCT	130
Db	478	GT	CTTTGTGCTCCAAAGAGGAGCGTCGGAACATT-TAAAGCCGTGGTTCCCTCCCTATCATGT	536
QY	131	ACT	CCCTTGATCTGTGCTCTGGCCCTCTGGGGAATATTCTGGTGGTGATCACCCTTGGCTT	190
Db	537	ACT	CCATCATTTTGTTCGTGGGCCCTACGTGGGCAATAGGGCTGGTGTGTGACCTATATCT	596
QY	191	TTT	TATAAGAGCCAGGCTATGACAGACGCTATCTTCTTGAACATGGCCATTGCAGACA	250
Db	597	ATT	TCAAGAGGCTCAAGACCATGACCGATACCTACCTGCTCAACTGGCGGTGCAGACA	656
QY	251	TCT	CTTTGTTCTTACTCTCCATTCTGGGAGTGAGTCATGCCACTGGTCGTGGGTTT	310
Db	657	TCT	CTTCTCCTGACCCCTTCCCTTCTGGGCTCAGCGCGGCCA---AGTCCTGGGTCT	713
QY	311	TC	AGCAATGCCAGTGCAGATTGCTTAAAGGCATCTATGCCATCAACTTTAACTGCGGGA	370
Db	714	TGG	GTGCACATTTTGCAGGCTCATCTTTGGCCATCTACAAGATGAGCTTCTTCAGTGSCA	773
QY	371	TGT	GTGCTGTACTTTGCAATTAGCATGGACCGGTACATCGCATCTTACAGCGGACTAAGT	430
Db	774	TG	CTCTACTTCTTTGCAATCAGCATTGACCGCTAGCTGGCCATCGTCGAGGCTGTCTCAG	833
QY	431	CAT	TCGGGTCGATCCAGAACACTACGCGCAGCAAAATCATCTGCCCTTGTGTGTGGG	490
Db	834	CT	CACGCCACCGTGCCCGGTCCTTCTCATCAGCAAGCTGTCCTGTGTGGGCATCTGGA	893
QY	491	GG	CTGCATCATCATCTCCAGCTCAACTTTTGTCTTCAACCAAAATPACAACCCCAAG	550
Db	894	TAC	TAGCCACAGTGTCTCCATCCCAAGAGCTCCTGTACAGTGAACCTCCAGAGGA-----	947
QY	551	GAG	CGAGTCTCTGTAACCCCAAGTAGCAAACTGTCTCCGAGCCCATCAGTGGGAAGTGC	610
Db	948	GC	AGCAGTGAGCAACGGATCGGATGCTCTCTCATCACAAGACATGTGGAGGCCCTTATCA	1007
QY	611	TGA	TGTGGGGCTTGAGCTACTCTTTGGTTCTTTATCCCTTTTGATGTTCAATGATATTTT	670

Db	1008	CCATCCAGGTGGCCAGATGGTGATCGGCTTTCTGGTCCCCCTGCTGCCCATGAGCTTCT	1067
Qy	671	GTTACACGTTCTATTGTCAAAACCTCTGGTGCAAGCTCAGAAATTTCTAAAAGGCACAAAGCCA	730
Db	1068	GTTACCTTGTTCATCATCCGACCCCTGCTCCAGGCAGCAACTTTGAGCGCAACAAGGCCA	1127
Qy	731	TCCGTGTAATCATAGCTGTGGTGTGTTGTTCTTGGCTTGTGCAGATTCCTCATACA---	787
Db	1128	TCAAGGTGATCATCGCTGTGGTGTCTTCATAGTCTTCCAGCTGCCCTACAATGGGG	1187
Qy	788	TGGTCTCTGTTGTGACGGCTGCTAATTTGGGTAAATGAACCGATCTCCGACAGCGAAA	847
Db	1188	TGGTCTCTGGCCAGACGGTGGCCAACTTCAACATCACAGTAGACACCTGTGAGCTCAGTA	1247
Qy	848	AGGTAATTTGGCTATACGAAACCTGTACAGAAAGTCTCTGGCTTTCTTGCACCTGCTCCCTGA	907
Db	1248	AGCAACTCAACATCGCCTACGACGTCACCTACAGCTGGCCTGCGTCCGCTGCTCGGTCA	1307
Qy	908	ACCTGTGCTCTACGCTTTTATTGGCGCAAGTTCAGAAACTACTTCTGAAGATCTTGA	967
Db	1308	ACCTTTCTTGTGACGGCTTCATCGGGGTCAAGTTCGCAACGATCTCTTCAAGCTCTTCA	1367
Qy	968	AGGACCTGTGGTGTGTGAGAAGGAAGTA	995
Db	1368	AGGACCTGGGCTGCCTCAGCCAGGASCA	1395
RESULT	12		
LOCUS	AR107241	1900 bp DNA	PAT 14-FEB-2001
DEFINITION	Sequence 18 from patent US 6107475.		
ACCESSION	AR107241		
VERSION	AR107241.1	GI:12821771	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1900)		
AUTHORS	Godiska, R., Gray, P. W. and Schweickart, V. Louise.		
TITLE	Seven transmembrane receptors		
JOURNAL	Patent: US 6107475-A 18 22-AUG-2000;		
FEATURES	Location/Qualifiers		
source	1..1900		
	/organism="unknown"		
BASE COUNT	431 a 593 c 459 g		417 t
ORIGIN			

Query Match	19.8%	Score 222;	DB 9;	Length 1900;
Best Local Similarity	55.5%	Pred. No. 4.1e-53;		
Matches 515;	Conservative	0;	Mismatches 400;	Indels 13;
				Gaps 4;

Qy	71	GTATTGTCGTCCTCTCAGGAGGTCAGGCAGTTCCTCCAGGCTATTTTGTACCGAATTCCT	130
Db	243	GTCTTTGTGTCCTCAAGAAGGCGTCGGGAACTT-FAAAGCCTGGTTCCCTCCCTATCATGT	301
Qy	131	ACTCCCTGATCTGTCTTTGGCTCCTCGGGAAATATCTCGTGGTGATACACCTTTGCTT	190
Db	302	ACTCCATCATTTGTTCTGGGGCTACTGGGCAATGGGCTGGTCTGTTTGACCTATATCT	361
Qy	191	TTTATAAGAGGCCAGGCTATGACAGAGCTCTATCTTTGAACATGGCCATTCGAGACA	250
Db	362	ATTTCAGAGGGTCAAGAGCATGACCGATACCTGCTCAACCTGGCGGTGGCAGACA	421
Qy	251	TCCCTCTTGTCTACTCTCCCAATCTGGCAGTGAGTCATCCCACTGGTGGTGGGTTT	310
Db	422	TCCCTCTCCTCTGACCCCTTCCCTCTGGGCCCTACAGCGGGCCA---AGTCTGGGCT	478
Qy	311	TCAGCAATGCCAGTCGCAAGTTGCTAAAAGCATCTATGCCATCAACTTTTAACTCGGGA	370
Db	479	TCGGGTGCACATTTTTCGAAGCTCATCTTTGGCATCTTACAAGATGAGCTCTTTCAGTGCA	538



[illegible]

QY	491	GGCTGTCACTCATCATCTCCAGCTCAACTTTTGTCTTCAACCAAAAATACAACACCCCAAG	550
DB	617	TACTAGCCACAGTGCCTCCATCCACAGAGCTCTGTGTACAGTACCTCCAGAGGA	670
QY	551	GCAGCGATCTCTGTGAACCCAACTACCAAGTACCAAACTGTCTCGGAGCCCATCAGTGGNAAGTGC	610
DB	671	GCAGCAGTGAAGAAGCATGCGATGCTCTCTCATCACAGAGCATGTGGAGGCGCTTTATCA	730
QY	611	TGATGTTGGGCGTTGAGCTACTCTTTGGTCTTTTATCCCTTTTGATCTTCATGATATTTT	670
DB	731	CCATCCAGTGGCCAGATGGTATCGGCTTCTGTGTCCTGCTGGCCATGAGCTTCT	790
QY	671	GTTACAGTGTTCATTGTCAAAACCTTGGTCAAGCTCAGAAATCTTAAAGGCACAAAGCCA	730
DB	791	GTTACTTGTCTCATCCGCACTCTCTCAGCAGCAAGCACTTTGAGCGCAACAAGGCCA	850
QY	731	TCCGTGTAATCATAGTGTGGTGTGTTCTGCTTGTCTGCTTGTCTAGATTCCTCATAACA	787
DB	851	TCAAGGTGATCATCGCTGTGGTGTCTTCATAGTCTTCCAGCTGCCCTACAATGGGG	910
QY	788	TGCTCTGCTTGTGAGGCGTGTAAATTTGGTGGTAAAAATGAACCGCATCTCCGACAGCGAAA	847
DB	911	TGCTCTGCGCCAGAGCGTGGCACTTCAACATCACAGTAGACCTGTGAGCTCAGTA	970
QY	848	AGCTAATTGGCTATACGAAAACTGTACAGAAGTCTCGCTTTCTGCTGCACTGCTGCCTGA	907
DB	971	AGCAACTCAACATCGCTACGAGCTCACCTACAGCCTGGCTGCGTGGCTGCTGCGTCA	1030
QY	908	ACCTGTGCTCTACGCTTTTATTTGGCAGAAGTTTCAAGAATCTACTTTCTGAAGATCTTGA	967
DB	1031	ACCTCTTCTGTGACGCTTTCATCGGCGTCAAGTTCCGCAACGATCTCTTCAAGCTCTTCA	1090
QY	968	AGGACCTGGTGTGTGAGGAAGGA	995
DB	1091	AGGACCTGGGCTGCCTACGCCAGGAGCA	1118
RESULT	15		
LOCUS	AR107239	2160 bp	DNA
DEFINITION	Sequence 14 from patent US 6107475.		PAT
ACCESSION	AR107239		
VERSION	AR107239.1	GI:12821769	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 2160)		
TITLE	Godiska, R., Gray, P.W. and Schweickart, V. Louise.		
JOURNAL	Seven transmembrane receptors		
FEATURES	Patent: US 6107475-A 14 22-AUG-2000;		
source	Location/Qualifiers		
	1. .2160		
	/organism="unknown"		
BASE COUNT	490 a. 646 c 545 g		479 t
ORIGIN			

[illegible]

Db 317 ATTCAAGAGGCTAAGACCATGACCGATACCTACCTGCTCAACCTGGCGGTGGCAGACA 376  
QY 251 TCCCTCTTTGTTCTTACTCTCCCAATTCGGGCACTGAGTCATGCCCACTGCTGGTGGGTTT 310  
Db 377 TCCCTCTCTCTGACCTTCCCTTCTGGGCTACAGCGCGGCCA---AGTCTGGGTCT 433  
QY 311 TCAGCAATGCCACGTGCAAGTTGCTAAAGGCACTATGCCATCAACTTTAACTGGGGGA 370  
Db 434 TCGGTGTCCACTTTTTCGAAGCTCATCTTTGCCATCTACAAGATGAGCTTCTTCAGTGGCA 493  
QY 371 TGCTGCTCCTGACTTGCATTAGCATGGACCGGTACATCGCCATTGTACAGGCGACTAAGT 430  
Db 494 TGCTCCTACTCTTTGCAATGACATGACCGGTACGTGCGCATCGTCCAGGTGCTCTCAG 553  
QY 431 CATTCGGGCTCCGATCCAGAACACTACGCGCAGCAAAATCATCTGCCCTTGTGTGGG 490  
Db 554 CTCACCGCACCGTGCCTGGGCTTCTCATCAGCAAGCTGTCTGTGTGGGCATCTGGA 613  
QY 491 GGCTGTAGTCATATCTCCAGCTCAACTTTTGTCTTCAACCAAAAATACACACCCCAAG 550  
Db 614 TACTAGCCACAGTGTCTCTCCATCCAGAGCTCTGTACAGTGACCTCCAGAGGA----- 667  
QY 551 GCAGGAGTGTCTGTAACCCCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGAAGCTGC 610  
Db 668 GCAGCAGTGAGCAACCGATGCCATGCTCTCTCATCACAGAGCATGTGGAGGCCCTTATCA 727  
QY 611 TGATGTTGGGCTTGAGTACTCTTTGGTTTCTTTATCCCTTTGATGTTTCATGATATTTT 670  
Db 728 CCATCCAGGTGGCCAGATGATGATCGGCTTTCTGTCTCCCTGCTGGCCATGAGCTTCT 787  
QY 671 GTTACACGTTCAATGTCAAAACCTTTGGTGCAAGCTCAGAAATTTCTAAAGGCACAAAGCCA 730  
Db 788 GTTACCTTGTATCATCTCCGCAACCTGTCTCCAGGCACGCAACTTTGAGCGCAACAAGGCCA 847  
QY 731 TCCGTTGTAATCATAGCTGTGGTGTCTTGTCTTGTGCTTGTGAGATTCTCTCATATAACA--- 787  
Db 848 TCAAGGTATCATCTGTGTGCTGGTCTTCTATAGTCTTCCAGCTGCCCTACATGGGG 907  
QY 788 TGGTCTCTGTTGTAGCGCTGCTAATTTGGGTAATGAACCGATCCTGCCAGAGCGAAA 847  
Db 908 TGGTCTTGCCCGCAGACGGTGGGCCAACTTCAACATCACCAAGTAGCACCTGTGAGCTCAGTA 967  
QY 848 AGCTAATTTGGCTATACGAAACTGTACAGAAAGTCTCTGGCTTTCTGCACTGCTGCCTGA 907  
Db 968 AGCAACTCAACATCGCTACGACGTCACCTACAGCTGGCCTGGCTGCCGTGCTGCCTCA 1027  
QY 908 ACCCTGTGCTCTAGCGCTTTTATTGGGCAAGTTCAGAAACTACTTTCTGAAGATCTTGA 967  
Db 1028 ACCCTTTCTGTACGCCCTTCTATCGCGGTCAAGTTCCGCCAAGCATCTCTTCAAGCTCTTCA 1087  
QY 968 AGGACCTGTGTGTGAGAGGAAGTA 995  
Db 1088 AGGACCTGGCTGCCTCAGCCAGGAGCA 1115

Search completed: May 24, 2001, 08:25:21  
Job time: 5532 sec

PR 11-OCT-1996; 96US-0028329.

XX (SCHE ) SCHERING CORP.  
XX Gish KC, Schall TJ, Vicari A, Wang W, Zlotnik A;  
XX WPI: 1998-101054/09.  
XX P-PSDB; W48086.  
XX Novel chemokines, e.g. thymus expressed chemokine - used for  
XX treating inflammatory conditions including asthma.  
XX Claim 5; Page 92-94; 202pp; English.  
XX The present sequence encodes a human dendritic cell chemokine receptor.  
XX Antibodies which bind to the protein can be used in detecting or  
XX diagnosing various immunological conditions related to expression  
XX of the protein. The nucleic acid can be used for screening and  
XX isolating DNA clones for the chemokines, especially from other  
XX species. The chemokine can be used in the treatment of conditions  
XX associated with abnormal physiology or development, including  
XX inflammatory conditions such as asthma.  
SQ Sequence 1119 BP; 261 A; 268 C; 257 G; 332 T; 1 other;  
  
Query Match 99.9%; Score 1118; DB 19; Length 1119;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1118; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ATGTTTTCAGCTCCAGTCAAGATATTTTGTCTCAGTCAATCTTCAATATTACTCAGTTG 60  
DB 1 atgttttcagctccagtgaaagtattttgtctcagtcgaataacttcaattactcagttg 60  
  
QY 61 ATCTGAGATGTTACTGTCTCTTCGAGGAGGTGAGGAGTCTTCCAGGCTATTGTGATC 120  
DB 61 attctgagatgttactgtctcttcgagtgagtcagtcagttctccaggtatttggatc 120  
  
QY 121 CGAATTGCTACTCTTGTATCTGTGTTCTTGGCCCTCTGGGGAATATCTGGTGGTATC 180  
DB 121 cgaattgctactcctgtatctgttcttggccctctggggaatatcttgggtggtatc 180  
  
QY 181 ACCTTTGCTTTTATAAGGCCAGGTCTATGACAGACGTCTATCTCTTGAACATGGCC 240  
DB 181 accttggcttttataagaagccaggtctatgacagacgtctatctcttgaaatggcc 240  
  
QY 241 ATTCAGACATCTCTTGTCTTACTCTCCCATCTGGGAGTGTGAGTATGCCACTGTG 300  
DB 241 attgcagacatctcttgttcttactctcccatctgggagtgagtcagtcactggt 300  
  
QY 301 GCGTGGGTTTTCAGCAATGCCACGTGCAAGTCTTAAAGGCATCTATGCCATCACTTT 360  
DB 301 gcgtgggttttcagcaatgccacgtgcaagtgtctaaagcagtcactatgccaacttt 360  
  
QY 361 AACTCGGGATGCTCTCTGACTTGTGATTAAGTACATGGACCGGTACATGCCATTTGTACAG 420  
DB 361 aactcgggatgctctctgacttgtgattagcatggacccggtacatgccattgtacag 420  
  
QY 421 GCGACTAAGTCATTCGGGCTCCGATCCAGAACACTACCCGCGAGCAAAATCATCTGCC 480  
DB 421 gcgactaagtcattccggctccgataccagaacactaccgcgagcaaaatcatctgcct 480  
  
QY 481 GTTGTGTGGGGCTGTAGTCATATCTTCAGCTCAACTTTTGTCTTCAACCAAAATAC 540  
DB 481 gttgtgtggggctgtagtcataatcttcagctcaacttttgtcttcaaccaaataac 540  
  
QY 541 AACACCCAGGAGGAGTCTGTGAACCCAGTACCAAACTGTCTCGAGGCCATCAGG 600  
DB 541 aacacccaggagagtgctgtgaacccagtaccaaactgtctcggagcccatcagg 600  
  
QY 601 TGGAGCTGCTGATCTTGGGCTTCCAGTACTCTTGTGTTTCTTTATCCCTTTGATGTC 660  
DB 601 tggagctgctgatcttgggcttccagtcactcttggttttctttatccctttgatgttc 660

661 ATGATATTTTGTACAGCTTCATTTCAAAACCTTGGTGAAGCTCAGAAATCTAAAGG 720  
661 atgatattttgttacagcttcatttcaaaaccttgggtgaagctcagaattctaaagg 720  
721 CACAAAGCCATCCGTGTAATCATAGCTGTGGTGTGTTTCTGGCTTGTGAGATTCT 780  
721 cacaaagccatccgtgtaatacatagctgtgggtgtgtttctggcttgcagattctc 780  
781 CATAACATGCTGCTGTGTGACGGCTGCTAAATTTGGTGAATTAACCCGATCTGCCAG 840  
781 cataacatgctgctgtgtgacggctgctaaatttgggtgtaataaacccgatcctgccag 840  
841 AGCGAAAGCTAATTCGCTATACGAAACTGTACAGAGTCTCGCTTCTTCTGCCTGTC 900  
841 agcgaaagctaatctcgctatacgaactgtcacagaagtcctggcttctcactgc 900  
901 TGCTGAACCTGCTGTCTACGCTTTTATTTGGGAGAGTTCAGAAACTACTTTCTGAAG 960  
901 tgctgaacctgctgtctacgcttttatttgggagagttcagaactacttctgaag 960  
961 ATCTTGAAGACCTGTGTGTGAGAGGAGTACAGTCTCAGGCTTCTCTGTGTC 1020  
961 atcttgaagacctgtgtgtgagaggaagtacagtcctcaggtcttctcctgtgcc 1020  
1021 GGGAGTACTCAGAAACATTTCTCGCAGACCACTGAGACCGCAGATAACGACAATGCG 1080  
1021 gggagtactcagaaacatttctcggcagaccagtgtgagaccgagataacgacaatg 1080  
1081 TCGTCTTCTACTATGTATAGAAAGTCTGCTCCCTAA 1119  
1081 tgcgtcttctactatgtatagaaagtctgctccctaa 1119  
1081 tgcgtcttctactatgtatagaaagtctgctccctaa 1119

RESULT 2  
A51971  
ID A51971 standard; cDNA; 1098 BP.  
XX A51971;  
XX 04-DEC-2000 (first entry)  
XX Primate (human) chemokine receptor CCR6 coding sequence.  
XX MIP-3-alpha; chemokine; CCR6 receptor; ligand: cytostatic; agonist;  
XX antagonist; anti-psoriatic; dermatological; immunosuppressive;  
XX anti-inflammatory; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
CDS 1..1098  
FT /\*tag= a  
FT /transl\_except= (pos:577..579, aa:Xaa)  
FT /note= "Xaa is not defined"  
XX /product= CCR6  
PN WO200046248-A1.  
XX 10-AUG-2000.  
XX 02-FEB-2000; 2000WO-US00511.  
XX 03-FEB-1999; 99US-0244281.  
XX (SCHE ) SCHERING CORP.  
XX Oldham ER, Honey B, Dieu-Nosjean M, Caux C, Zlotnik A;  
XX WPI: 2000-543477/49.  
XX P-PSDB; Y97077.  
XX Novel methods for modulating the migration of cells within or to the  
XX skin using MIP-3-alpha or CCR6 agonists or antagonists, useful for

PT treating skin disorders, e.g. cancer  
XX Disclosure; Page 51-53; 61pp; English.  
XX  
CC The MIP-3-alpha chemokine is expressed in inflamed skin cells. This  
CC chemokine is the ligand for the CCR6 receptor. MIP-3-alpha or CCR6  
CC agonists or antagonists can be used to modulate the migration of a cell  
CC within or to the skin of a mammal. MIP-3-alpha can be used to purify a  
CC population of cells expressing a MIP-3-alpha receptor. The methods are  
CC useful for treating a mammalian subject with a skin disease or condition,  
CC e.g. cancer, metastasis, psoriasis, atopic or contact dermatitis,  
CC systemic lupus erythematosus and lichen ruber planus, or for treating  
CC skin transplants or grafts.  
XX  
SQ Sequence 1098 BP; 254 A; 263 C; 253 G; 327 T; 1 other;

Query Match 98.0%; Score 1097; DB 21; Length 1098;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1097; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTTTCGACTCCAGTGAAGATTATTTGTGTCAGTCAATATCTTCAATATCTCACTG 60  
DB 1 atgttttcgactccagtgaaatttttgttcagtcataacttcattactcagttg 60  
QY 61 ATCTGAGATGTTACTGCTGCTGCTGCAGGAGGTCAGGCGAGTTCTCCAGGCTATTCTAC 120  
DB 61 attctgagatgttacttgctgcttcgagaggtcaggtcaggttcctccaggtattgtac 120  
QY 121 CGAATTCCTTACTCCTTGATCTGTTGCTTCTGGGCTCCTCGGGGAATATCTGGTGGTGATC 180  
DB 121 cgaatgctactccttgatctgttcttctggcctcctcgtgggaattctcgtgtggtatc 180  
QY 181 ACCTTTGCTTTTATAGAGGCCAGTCTATGACAGAGCTCTATCTTGAACATGGCC 240  
DB 181 acctttgcttttataagaagccaggtctatgacagcgtctatctcttgaacatggcc 240  
QY 241 ATTGCAGACATCTCTTTGTTCTTACTCTCCCATCTCGGCGAGTGCATGCCACATGGT 300  
DB 241 attgcagacatctcttcttcttactctcccatctcggcagtgagtcattgcccactcgt 300  
QY 301 CGGTGGGTTTTCAGCAATGCCAGTGCAGTGTCTAAAGGCATCTATGCCATCAACTTT 360  
DB 301 gcgtgggttttcagcaatgccagtcgaagtgtcctaagaagcattctatgccaacttt 360  
QY 361 AACTGGGGATGCTCTCTGACTTGCATTAGCATGACCGGTACATCGCCATTGTACAG 420  
DB 361 aactgvggatgctctctgacttgacttgacattagcatgacccggtacatcgccattgtacag 420  
QY 421 GGACATAAGTCAATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCCTT 480  
DB 421 gcgactaaagtcattccggctcccgatccagaaactaccgacgagcaaaatcatctgcctt 480  
QY 481 GTTGTGTGGGGTGTCAGTCAATCTCCAGTCAACTTTTGTCTTCAACCAAAAATAC 540  
DB 481 gttgtgtggggctgcagtcacatctccagtcgaactttgtcttcaacccaaaatac 540  
QY 541 AACACCCAGGACGATGCTGTGAACCCCAAGTACAAACTGTCTCGGAGCCCATCAGG 600  
DB 541 aacacccaggcagatgctgtgaaccccaagtaccaanactgtctcggagcccatcagg 600  
QY 601 TGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGGTTTCTTATCCCTTTGATCTTC 660  
DB 601 tgaagctgctgagtgtggggcttgagctactcttggtttcttataccctttgatcttc 660  
QY 661 ATGATATTTTGTACAGTTCATTGTGCAAAACCTTGGTCAAGCTCAGAATCTTAAAGG 720  
DB 661 atgatattttgtacagttcattgtgcaaaaccttggtgcaagctcagaattctaaagg 720  
QY 721 CACAAGGCCATCCGTGTATCATPATGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
DB 721 cacaagccatccgtgtatcatpatgtgtgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780

QY 781 CATAACATGGTCTCTCTTGTGACGGCTGCTAATTTGGGTAAATGAACCGATCCTGCCCAG 840  
DB 781 cataacatggctctcttgtgacggctgctaatTTGGGTAAATGAACCGATCCTGCCCAG 840  
QY 841 AGCGAAAGCTTAATTGGCTATACGAAACTGTACAGAGTCTCTGGCTTTCTGCACTGC 900  
DB 841 agcgaagagcttaattggctatacgaactgtcacagaagctcgtgcttctcactgc 900  
QY 901 TGCCTGAACCTGCTGCTCTACGCTTTTATTTGGGCAAGTTCAGAACTACTTCTGAAG 960  
DB 901 tgcctgaacctgctgctctacgcttttatttggcagaagttcagaactacttctgaag 960  
QY 961 ATCTTGAAGGACCTGCTGGTGTGTGAGAGGAAGTACAAGTCTCAGGCTTCTCCTGTGC 1020  
DB 961 atcttgaaggacctgctgggtgtgtgagaggaagtaacaagtcctcaggtcttctctgtgc 1020  
QY 1021 GGGAGGTACTCAGAAACATTTCTCGGACACACAGTGCAGACCGCAGATACCAACAATGG 1080  
DB 1021 gggaggtactcagaaacatttctcggcagaccagtgagaccgcagataacgacaatgcg 1080  
QY 1081 TCGTCTTCACTATGTGA 1098  
DB 1081 tcgtcttcactatgtga 1098  
RESULT 3  
A30632 ID A30632 standard; cDNA; 1137 BP.  
XX AC A30632;  
XX XX 21-AUG-2000 (first entry)  
XX DE Human G protein-coupled receptor EB11 cDNA.  
XX KW G protein-coupled receptor; GPCR; constitutively active;  
KW intracellular loop 3; transmembrane domain 6; drug screening;  
KW agonist; antagonist; ss.  
XX OS Homo sapiens.  
XX PN WO200022129-A1.  
XX PD 20-APR-2000.  
XX PF 12-OCT-1999; 99WO-US23938.  
XX PR 13-OCT-1998; 98US-0170496.  
XX PA (AREN-) ARENA PHARM INC.  
XX PI Behan DP, Chalmers DT, Liaw CW;  
XX WPI; 2000-329165/28.  
XX P-PSDB; Y90629.  
XX PT Non-endogenous constitutively activated human G protein-coupled  
PT receptors, useful for identifying agonists for use as pharmaceutical  
PT agents -  
XX PS Example 1; Page 157; 341pp; English.  
XX CC The invention relates to constitutively active, non-endogenous versions  
CC of endogenous human orphan G protein-coupled receptors (GPCRs, Y90643-  
CC Y90677 and Y90683-Y90687), and to DNA encoding them (A30709-A30743 and  
CC A30775-A30779). The mutant proteins of the invention contain a  
CC mutation in a portion of the protein comprising intracellular loop 3  
CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, x,  
CC is substituted for an endogenous residue in IC3 at a position 16 amino  
CC acids N-terminal of an endogenous proline in TM6 to form a sequence  
CC x-(AA)15-pro. The endogenous amino acid is selected from Lys, His, Arg or  
CC Ala, and is preferably Lys. When the endogenous residue at this position  
CC is Lys, this residue is replaced by His, Arg or preferably Ala. The 15



Query Match 19.8%; Score 222; DB 15; Length 1900;  
Best Local Similarity 55.5%; Pred. No. 5.4e-60;  
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

```
QY 71 GTTACTGCTCTTCAGAGGTGACGAGTCTCCAGGCTATTGTACCGAATTGCCT 130
   || ||||| || || || || || || || || || || || || || || || ||
Db 243 gtctttgtctcaagaagagcgtgcgaactt-taaagcctggttccctccctcatgt 301

QY 131 ACTCCTTGATCTGTCTTGCTGCTCCTCGGGAATATTCTGGTGGTGATCACCTTTGCTT 190
   || ||||| || || || || || || || || || || || || || || || ||
Db 302 actccatcatctgttctcgtggcctactggcaatggctggctgttgacatcatct 361

QY 191 TTTATAAGAGGCGAGGTCTATGACAGAGCTCTATCTTTGAACATGCGCCATTGCAGACA 250
   || ||||| || || || || || || || || || || || || || || || ||
Db 362 attcaagagctcaagaccatgacgatcactacgtctcaacctggcgtggagagaca 421

QY 251 TCCTCTTTGTTCTTACTCTCCCATTTGGCGAGTGTGATGATGCTGCTGGGTTT 310
   || ||||| || || || || || || || || || || || || || || || ||
Db 422 tctcttctctgacccctccctctctggcctacagcggcca---agtcctgggtct 478

QY 311 TCAGCAATGCCAGCTGACAGTTTCTTAAAGGCACTCTATGCCATCACTTTAACTGCGGGA 370
   || ||||| || || || || || || || || || || || || || || || ||
Db 479 tgggtgtccacttttgcaagcctcatcttggccatccacaagatgagctcttcagtggca 538

QY 371 TGCTGCTCTGACTTGCATTAGCATGGACCGGTACATCGCCATTGTACAGCGGACTAAAGT 430
   || || || || || || || || || || || || || || || || || || || ||
Db 539 tgctctactcttctgcatcagcatgaccgtacgtacgtggccatgctccaggtgctcag 598

QY 431 CATTCGGGCTCGAYCCAGAACACTACCGCGCAGCAAAATCATCTGCTTGTGTGGG 490
   || || || || || || || || || || || || || || || || || || || ||
Db 599 ctacccgacccgtgcccgtctctctcatcagcaagctgtctgtgtgggcattcga 658

QY 491 GCGTGTGAGTCATCTCCAGCTCAACTTTGTCTTCAACCAAAATACACACCCCAAG 550
   || || || || || || || || || || || || || || || || || || || ||
Db 659 tactagccacagtgtctccatcccccagagctcctgtcacagtgaacctccagagga----- 712

QY 551 GCAGCATGCTCTGAAACCAAGTACCAAACTGCTCGGAGCCCATCAGTGGGAAGCTGC 610
   || || || || || || || || || || || || || || || || || || || ||
Db 713 gcagcagtgaagcagcagatgcgatctctctcatcagacagcagatgtggagcccttatca 772

QY 611 TGATGTGGGCTTGAGCTACTCTTTGGTTCCTTTTTCCTTTGATGTTTCATGATATTTT 670
   || || || || || || || || || || || || || || || || || || || ||
Db 773 ccatccaggtggccagatgggatcggtcttctgtgctccctcgtggccatgagcttct 832

QY 671 GTTACAGCTTCTATGTCAAACCTTGGTGCAGCTCAGATTCATAAGGCACAAAGCCA 730
   || || || || || || || || || || || || || || || || || || || ||
Db 833 gttacctgtcatctccgacccctgctccagcagcgaacttggagcgaacaaggcca 892

QY 731 TCCGTGTAATCATAGTGTGGTGTGTGTCTGCTGCTGTCAGATTCCTCATAAACA--- 787
   || || || || || || || || || || || || || || || || || || || ||
Db 893 tcaagtgatcatcgtgtggtcgtgtgtcttcatagcttccagctgccttcaaatgggg 952

QY 788 TGTCTCTGTTGTGAGGCTGTCTAATTTGGGTAAATGAAACCGATCTCCAGAGCGCAA 847
   || || || || || || || || || || || || || || || || || || || ||
Db 953 tggctcgtggccagacggtggcacaattcaacatccaccagtagcacctgtgagctcagta 1012

QY 848 AGCTAATGGCTATACGAAACTGTACAGAGTCTCTGCTGCTGCTGCTGCTGA 907
   || || || || || || || || || || || || || || || || || || || ||
Db 1013 agcaactcaacatccctacagcgtcacctacagcgtggcgtgctgctgctgctga 1072

QY 908 ACCCTGTGCTCTACGCTTTTATTTGGCAGAAGTTTCAAGAACTACTTCTTGAAGATCTTGA 967
   || || || || || || || || || || || || || || || || || || || ||
Db 1073 acccttctgtacgccttcacgtggcgtcaagtctccgcaacgatctcttcaagctcttca 1132

QY 968 AGGACCTGTGGTGTGAGAGGAAGTA 995
   || || || || || || || || || || || || || || || || || || || ||
Db 1133 aggaacctgggtgcctcagccaggagca 1160
```

RESULT 5  
V18349  
ID V18349 standard; DNA; 1900 BP.  
XX

```
AC V18349;
XX XX 25-SEP-1998 (first entry)
XX
DE Human V31 seven transmembrane (7TM) receptor genomic DNA exon 3.
XX V28; placenta; seven transmembrane receptor; 7TM; signal transduction;
KW immunology; inflammation; V31; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT intron 1..168 /*tag= a
FT exon /*tag= b
FT /*tag= number= Exon 3
FT /*tag= 169..1245
FT /*tag= 169..1245
FT /*tag= c
FT 3'UTR 1246..1900 /*tag= d
FT
XX US5759804-A.
PN
XX 02-JUN-1998.
PD
XX 17-NOV-1993; 93US-0153848.
PF
XX 17-NOV-1992; 92US-0977452.
PR
XX (ICOS-) ICOS CORP.
PA
PI Godiska R, Gray PW, Schweickart VL;
XX WPI; 1998-332132/29.
DR
XX DNA encoding V28 seven transmembrane receptor polypeptide - useful
PT for producing recombinant polypeptide and anti-V28 antibodies, and
PT in screening assays for V28 agonists and antagonists
XX
XX Example 3; Columns 43-48; 56pp; English.
XX
CC The present sequence represents the third exon of the V31 genomic
CC DNA (V18345) which was isolated from a human placental genomic library.
CC The third exon encodes for the last 358 amino acids of the V31 seven
CC transmembrane (7TM) receptor (W48724). The invention claims for the
CC full length V28 genomic DNA sequence isolated from a human placenta
CC genomic library. V28 (W48722) and V31 proteins are 7TM receptors which
CC are probably involved in signal transduction. The invention also claims
CC that cells transformed with V28 DNA can be used to produce the
CC recombinant polypeptide, to produce anti-V28 antibodies or in screening
CC assays for V28 agonists or antagonists. The antibodies, agonists and
CC antagonists could then be used to modulate V28 receptor-ligand binding,
CC for e.g. in immunological and/or inflammatory events in vivo.
XX
SQ Sequence 1900 BP; 431 A; 593 C; 459 G; 417 T; 0 other;
```

```
Query Match 19.8%; Score 222; DB 19; Length 1900;
Best Local Similarity 55.5%; Pred. No. 5.4e-60;
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

QY 71 GTTACTGCTCTTCAGAGGCTCAGGAGTCTCCAGGCTATTGTACCGAATTGCCT 130
   || ||||| || || || || || || || || || || || || || || || ||
Db 243 gtctttgtctccaaagaagacgtgcgaactt-taaagcctggttccctccctcatgt 301

QY 131 ACTCCTTGATCTGTCTTCTTGCTGCTCCTCGGGAATATTCTGGTGGTGATCACCTTTGCTT 190
   || ||||| || || || || || || || || || || || || || || || ||
Db 302 actccatcatctgttctcgtggcctactggcgaatggctggctgttgacatcatct 361

QY 191 TTTATAAGAGGCGAGGTCTATGACAGAGCTCTATCTTTGAACATGCGCCATTGCAGACA 250
   || ||||| || || || || || || || || || || || || || || || ||
Db 362 attcaagagctcaagaccatgacgatcactacgtcgaacctgtcgaacctggtggagagaca 421
```



```
QY 671 GTTACAGTTCATGTCATAAACCCTGGTGTGAAGCTCAGAAATCTAAAAAGGCACAAAGCCA 730
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 833 gttacctgtgtcatatccgcacacctgtccaggcaagcaactttgagcgcaacaggcca 892
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 731 TCCGTGTAAATCATAGCTGTGGTGTGTTGTTCTTGGCTTGTTCAGATTCTCTATACAA--- 787
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 893 tcaagggtgatacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 952
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 788 TGGTCTCTGTGTGACGGCTGCTAATTTGGGTAAATGAACCGATCCCTGCGCAGAGCGAAA 847
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 953 tggctctggccagacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 848 AGCTAATTTGGCTATACAAAACCTGTACAGAAAGTCTCTGGCTTTCTCTGACATGCTGCTGTA 907
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1013 agcaactcaacatcgctacagcgtcaccctacacgctggtgtgtgtgtgtgtgtgtgtgt 1072
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 908 ACCCTGTGCTCTACGCTTTTATTTGGGAGAGATTCAGAAACTACTTTCTGAAGATCTTGA 967
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1073 acctcttctgtacgcttcacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 968 AGGACCTGTGTGTGTGAGAAAGGAAGTA 995
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1133 aggaacctgggtgctcctcagccaggagca 1160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

## RESULT 7

```
Q66153
ID Q66153 standard; DNA; 2058 BP.
XX
AC Q66153;
XX
DT 02-FEB-1995 (first entry)
XX
DE Putative seven transmembrane receptor (V31) coding sequence.
XX
KW Primer; seven transmembrane receptor; receptor; amplification; PCR;
KW polymerase chain reaction; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 166..1398
FT FT /*tag= a
FT FT /product= V31 putative transmembrane receptor.
XX
PN W09412635-A.
XX
PD 09-JUN-1994.
XX
PF 17-NOV-1993; 93WO-US11153.
XX
PR 17-NOV-1992; 92US-0977452.
XX
PA (ICOS-) ICOS CORP.
XX
PI Godiska R, Gray PW, Schweickart VL;
XX
DR WPI; 1994-200264/24.
DR P-PSDB; R53743.
XX
PT DNA encoding seven transmembrane receptors - used to develop
PT prods. for use as therapeutic or diagnostic agents for conditions
PT involving the receptors.
XX
PS Example 2; Page 44-46; 100pp; English.
XX
CC Two primers (Q66148, Q66149) were used to amplify human genomic DNA
CC purified from leukocytes. Approximately 1000 clones were isolated
CC after the initial amplification reaction and probed with sequences
CC specific for seven transmembrane receptors IL8R1, AN2R and R20.
CC Clones which did not hybridise were then chosen for sequence
CC analysis. Three new clones were identified that appeared to encode
CC seven transmembrane receptor segments. Two more primers (Q66151,
```

```
CC Q66152) were used to isolate a full length version of one of these
CC clones (this sequence) which was designated V31.
XX
SQ Sequence 2058 BP; 472 A; 632 C; 504 G; 448 T; 2 other;
    Query Match 19.8%; Score 222; DB 15; Length 2058;
    Best Local Similarity 55.5%; Pred. No. 5.6e-60;
    Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;
QY 71 GTTACTGTCTCTCTGTCAGAGGTCAGGAGTTCCTCCAGGCTATTGTACCGAATTCGCT 130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 396 gcttttgtctcaagaagcagcgtgcggaactt-taaagcctgtgttctccctccatcatgt 454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 131 ACTCTCTGTATCTGTCTTGGCTCTCTGGGAATATTCTGGGTGTGATCACCTTTGCTT 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 455 actccatcatctgttctgttgcctactgtggaatggctggtcgtgtgtgtgtgtgtgtgt 514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 TTTATAAAGGAGCCAGGCTATGACAGAGCTGTATCTTGTACACATGGCCATTTGCAGACA 250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 attcaagagctcaagaccatgacgatacctacactgctcaacctggcgtggcagaca 574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 251 TCCTCTTTGCTTACTCTCCATTTCTGGGAGTGTGATGATGATGATGATGATGATGATGAT 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 575 tctcttctctgaccttctcttctggtggtcctacagcgcgcga---agtcctgggtct 631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 311 TCAGCAATGCCAGCTGCAAGTTGCTAAAGGATCTATGCCATCAACTTTAACTGCGGGA 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 632 tgggtgtccacttttgcaagctcatcttggccatctcaagatgagcttctcagtgcca 691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 371 TGCTGTCTCTGACTTGTGATGACGACCGGTGTACATCGCCATTTGTACAGCGACTAAGT 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 692 tgctctactcttctgcatcagcattgacgtacgtggccatcgctccagcgtgtctcag 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 431 CATTCGGCTTCGATCCAGAACACATACCGCGAGCAAAATCATCTGCTTGTGTGTGGG 490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 752 ctcaaccgcaacgtgcccgcgtctctctcatcagcagcgtgtctgtgtgtgtgtgtgtgt 811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 491 GGCTGTGATCATCATCTCTCAGCTCAACTTTGCTTCAACCAAAATACAAACCCCAAG 550
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 812 tactagccacagtgctctctccatcccccagagctcctgtacagtgacctccagagga 865
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 551 GCAGCAGTGTCTGTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGCTGGAAGCTGC 610
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 866 gcagcagtgagcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 925
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 611 TGATGTTGGGCTTGAGCTACTCTTTGGTGTCTTTATTCCTTTGATGTTGATGATATTT 670
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 926 ccataccagtggtccagatggtgagcgttcttctgtgtgtgtgtgtgtgtgtgtgtgt 985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 671 GTTACAGCTTCAATTGTCAAAACCTTGGTGTGCAAGCTCAGAAATTTAAAGGCACAAAGCCA 730
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 986 gttacctgtgtcatatccgcacacctgtccagcagcagcagcagcagcagcagcagcag 1045
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 731 TCCGTGTAAATCATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 787
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1046 tcaaggtgatacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 788 TGGTCTCTGTGTGACGGCTGTCTAATTTGGGTAAATGAACCGATCCCTGCGCAGAGCGAAA 847
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1106 tggctctggccagacagcgtgtggaacttcaacatcaccagtagcacctgtgagctcagta 1165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 848 AGCTAATTTGGCTATACGAAACTGTACAGAACTCTCGGCTTTTCTGCACTGTGCTCTGA 907
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1166 agcaactcaacatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 908 ACCCTGTGCTCTACGCTTTTATTTGGCAGAAATTCAGAACTACTTTCTCAAGATCTTGA 967
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1226 acctcttctgtacgcttctcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 968 AGGACCTGTGTGTGTGAGAAAGGAAGTA 995
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1286 aggaacctgggtgctcctcagccaggagca 1313
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY	251	TCCTCTTTGTTTACTCTCCCATTCCTGGCAGTGAATCATGCCACTGGTGGTGGTTT	310
Db	575	tcctctctctgacctctccctctggcctacagcgcgcca---agtcctgggtct	631
QY	311	TCAGCAATGCCACGTGCAAGTTGCTAAAAGGCATCTATGCCATCAACTTTAACTGCGGGA	370
Db	632	tcgggtgtccacttttgcaagctcatctttgcaatctacaagatgagcttcttcagtgcca	691
QY	371	TGCTGCTCCTGACTTGCATTAGCATGACGACCGGTACATCGCCATTGTACAGCGACTAAGT	430
Db	692	tgctctactcttctgcatcagcattgacctcaegtgcctggccatcgctccaggctgtctcag	751
QY	431	CATTTCGGGCTCCGATCCAGAACACATPACCGCGGAGCAAAATCATCTGCCTTGTGTGTGGG	490
Db	752	ctcagcgccacogtgccecgctcctctctcatcagcaagctgtcctgtgtgggcatctgga	811
QY	491	GGCTGTGAGTCATCATCTCCAGCTCAACTTTTGCTCTCAACCAAAAATCAACACCCCAAG	550
Db	812	tactagccacagtgctctccatcccgaagctcctgtacagtgacctccagagga-----	865
QY	551	GCAGCGATGCTGTGAACCCAAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGAAGCTGC	610
Db	866	gcagcagtgagcaagcagtgatgctctctcatcacagagcatgtggaggcccttlatca	925
QY	611	TGATGTTGGGCTTGAGCTACTCTTTTGGTTTCTTTFATCCCTTTTGATGTTTCATGATATTTT	670
Db	926	ccatccaggtggccagatgggatcggtcttctgtccccctgtggccatgagctctct	985
QY	671	GTTACAGGTTTCAATGTGCAAAACCTTGGTGCAAGCTCAGAAATCTAAAAGGCACAAAGCCA	730
Db	986	gttacctgtcatcatccgcacccctgctccaggcagcaactttgagcgcaacaggcca	1045
QY	731	TCGGTGTAAATAGCTGCTGTGCTTGTCTTTTGTGCTTTGTCCAGTTCCTTCATPAACA---	787
Db	1046	tcaagtgatcatcgctgtgctgtggtcttcatagcttccagctgccctacaatgggg	1105
QY	788	TGGTCTGCTTGTGACGGCTGCTAATTTGGGTAAATGAACCCATCCTGCCACAGCGAAA	847
Db	1106	tggtctggccagcaggtggccaaactcaacatcacagtagcacctgtgagctcagta	1165
QY	848	AGCTAATTGGCTATACGAAAACCTGTACAGAAAGTCTCCTGCTTTCCCTGCACTGCTGCTGA	907
Db	1166	agcaactcaacatgcctacgagtcacctaagcctgcctgcgtgcgtgcgtca	1225
QY	908	ACCCTGTGCTCTACGCTTTTATTGGSCAGAAAGTTCAGAAACTACTTTCTGAAGATCTTGA	967
Db	1226	acccttcttglacgcttcatcggtcgaagttccgcgaagatctcttcaagctcttca	1285
QY	968	AGGACCTGTGCTGTGGAAGAAGTA	995
Db	1286	aggacctggctgcctcagccaggagca	1313
RESULT	9		
AD600			
ID	A90600	standard; DNA; 2058 BP.	
AC	XX		
AC	A90600;		
XX			
DT	26-JAN-2001	(first entry)	
DE		Genomic clone of 7TM receptor V31.	
DE		Seven transmembrane receptor; 7TM; heptahelical; serpentine;	
KW		G-protein-coupled; v28; v31; v112; R20; R2; R12; RM3; gene therapy;	
KW		cancer; ss.	
XX			
OS		Homo sapiens.	
XX			
PN		US6107475-A.	
XX			
PD		22-AUG-2000.	
XX			



CC Clone PBMC75 was isolated and the V31 cDNA insert in the clone was  
XX designated cDNA V31-B (066160).

SQ Sequence 2160 BP; 490 A; 645 C; 546 G; 479 T; 0 other;

Query Match 19.8%; Score 222; DB 15; Length 2160;  
Best Local Similarity 55.5%; Pred. No. 5.8e-60;  
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

QY 71 GTTACTGTGCTCTTGCAGGAGGTCCAGGCTTCTCCAGGCTATTGTACCGAATTCGCT 130  
Db 198 gtctttgtctccaaagaagcgtgcggaactt-taaagcctggttctccctcatatcgt 256  
QY 131 ACTCCTTGATCTGTTCTTGGCCCTCTGGGGAATATTCTGGTGTGATCACTTTTGCTT 190  
Db 257 actccatattgttctggcctactggtgcaatggcgtgctgttgacctatct 316  
QY 191 TTTATAGAAGCCAGGCTATGACAGAGCTGTATCTTGAACATGGCCATTCAGACA 250  
Db 317 attcaagaggtcaagaccatgacctacactgctcaacctggtggcgagaca 376  
QY 251 TCGTCTTTGTTCTTACTCTCCATTTCTGGCAGTGTGAGTCACTGCTGGGT 310  
Db 377 tctcttctctgaccttccctctctgggctacagcgcgcca---agtcctgggtct 433  
QY 311 TCAGCAATGCCAGCTGCAAGTTGCTAAAGGCACTATGCTCCATCACTTTAACTGCGGA 370  
Db 434 tgggtgccacttttgcagctcatttggccatctcaagatgagcttctcagtgga 493  
QY 371 TCGTCTCTGACTGACTATGACATGAGCGGCTGATCGCATTCGACAGCGACTAAGT 430  
Db 494 tgcctctacttcttgcacagcatgaccgtcagtgccatggtgacctccaggtgtctcag 553  
QY 431 CATTCGGGCTCCGATCCAGAACTACCGCGCAGCAAAATCATCTGCCCTTGTGTGGG 490  
Db 554 ctccagccacgtgccgcgtctctctcatcagcaagctgctgtggtggcatctga 613  
QY 491 GCGTGTAGTATCATCTCCAGCTCAACTTTGTCTTCAACCAAAATACACACCCAG 550  
Db 614 tactagccacagtgtctccatccagagctctgtacagtgcacctccagagga----- 667  
QY 551 GCAGCGATGCTGTGAACCCCAAGTAGTACCAAACTGTCTCGAGCCCATCAGTGGAGCTGC 610  
Db 668 gtagcagtgagcaagcagatgogctctctctcatcagagcagcatgtggagggcctttatca 727  
QY 611 TGATCTGGGCTCTGAGTACTCTTTTGTGTTCTTATCCCTTGTATGTTTCATGATATTT 670  
Db 728 coactcagggtggccagatggtgatcggtcttctggtccctctggtgacctgacctct 787  
QY 671 GTTACAGTTTCAATGTCAAACTTGGTGAAGCTCAGATTTCTAAAGGCAAGACCA 730  
Db 788 gttactgttcatcatccgcacacctgctccagggcagcaactttgagcgcaacaaggcca 847  
QY 731 TCCGTGTAATCATAGCTGTGGTGTGCTTGTCTTCTGGCTGTGATTCCTCATAACA--- 787  
Db 848 tcaaggtgatactcgtgtggtgtgtctcatagttctccagctgacctcaaatggggg 907  
QY 788 TGGTCTGCTGTGAGGCTGTCTAATTTGGGTAAATGAACCGATCTGCCAGAGCGAAA 847  
Db 908 tggctctgggcccagacggtggcacaactcaacatcaccagtagcaccctgtgagctcagta 967  
QY 848 AGCTAATTTGGCTATACGAAACTGTACAGAAAGTCTCGCTTTCTCTGCACTGCTGCTGA 907  
Db 968 agcaactcaacatcgctacagcgtccacctacagcctggcctgctgctcgtcgtca 1027  
QY 908 ACCCTGTGCTTACGCTTTTATTTGGGAGAGAACTTTCAGAAACTTCTTGAAGATCTTGA 967  
Db 1028 acccttcttctgacgccttctcgcgtcaggttccgcaacgatctcttccagctcttca 1087  
QY 968 AGGACCTGTGGTGTGAGAGAGTA 995  
Db 1088 aggaacctgggtgctcagcgaggaca 1115

RESULT 11

VI18347

XX V18347 standard; cDNA; 2160 BP.

AC V18347;

DT 25-SEP-1998 (first entry)

XX Human V31 seven transmembrane (7TM) receptor cDNA.

XX V28; placenta; seven transmembrane receptor; 7TM; signal transduction; immunology; inflammation; V31; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 64..1200

FT /tag= a

XX /product= "Human V31 seven transmembrane receptor"

XX US5759804-A.

XX 02-JUN-1998.

XX 17-NOV-1993; 93US-0153848.

XX 17-NOV-1992; 92US-0977452.

XX (ICOS-) ICOS CORP.

XX Godiska R, Gray PW, Schweickart VL;

XX WPI; 1998-332132/29.

XX DR P-PSDB; W48724.

XX DNA encoding V28 seven transmembrane receptor polypeptide - useful for producing recombinant polypeptide and anti-V28 antibodies, and in screening assays for V28 agonists and antagonists

XX Example 3; Columns 35-40; 56pp; English.

XX The present novel sequence represents the V31 cDNA sequence isolated from the peripheral blood mononuclear cell cDNA library. The invention claims for the full length V28 genomic DNA sequence (V18343) isolated from a human placenta genomic library. V28 (W48722) and V31 (W48724) proteins are seven transmembrane (7TM) receptors which are probably involved in signal transduction. The invention also claims that cells transformed with V28 DNA can be used to produce the recombinant polypeptide, to produce anti-V28 antibodies or in screening assays for V28 agonists or antagonists. The antibodies, agonists and antagonists could then be used to modulate V28 receptor-ligand binding, for e.g. in immunological and/or inflammatory events in vivo.

XX Sequence 2160 BP; 490 A; 646 C; 545 G; 479 T; 0 other;

Query Match 19.8%; Score 222; DB 19; Length 2160;  
Best Local Similarity 55.5%; Pred. No. 5.8e-60;  
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

QY 71 GTTACTGTGCTCTTGCAGGAGGTCCAGGCTTCTCCAGGCTATTGTACCGAATTCGCT 130

Db 198 gtctttgtctccaaagaagcgtgcggaactt-taaagcctggttctccctcatatcgt 256

QY 131 ACTCCTTGATCTGTTCTTGGCCCTCTGGGGAATATTCTGGTGTGATCACTTTGCTT 190

Db 257 actccatcattgttctggtggcctactggtgcaatggcgtggtgtgtgacctatct 316

QY 191 TTTATAGAAGCCAGGCTGTATGACAGAGCTCTATCTTGAACATGGCCATTCAGACA 250

Db 317 attcaagaggtcgaagaccatgacctacactgctcaacctggtggcgagaca 376











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2001, 06:50:54 ; Search time 1188.46 seconds  
(without alignments)  
8225.514 Million cell updates/sec

Title: US-08-887-977-9  
Perfect score: 1119  
Sequence: 1 ATGTTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9623517 seqs, 4368049070 residues  
Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*  
44: gb\_est44:\*  
45: gb\_est45:\*  
46: gb\_est46:\*  
47: gb\_est47:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estin5:\*  
68: em\_estom1:\*  
69: em\_estom2:\*  
70: em\_estov1:\*  
71: em\_estov2:\*  
72: em\_estpl1:\*  
73: em\_estpl2:\*  
74: em\_estpl3:\*  
75: em\_estpl4:\*  
76: em\_estpl5:\*  
77: em\_estpl6:\*  
78: em\_estpl7:\*  
79: em\_estpl8:\*  
80: em\_estpl9:\*  
81: em\_estpl10:\*  
82: em\_estro1:\*  
83: em\_estro2:\*  
84: em\_estro3:\*  
85: em\_estro4:\*  
86: em\_estro5:\*  
87: em\_estro6:\*  
88: em\_estro7:\*  
89: em\_estro8:\*  
90: em\_estro9:\*  
91: em\_estro10:\*  
92: em\_estro11:\*  
93: em\_estro12:\*  
94: em\_estro13:\*  
95: em\_estro14:\*  
96: em\_estro15:\*  
97: em\_estro16:\*  
98: em\_estro17:\*  
99: em\_estro18:\*  
100: em\_estro19:\*  
101: em\_estro20:\*  
102: gb\_est25:\*  
103: gb\_est26:\*  
104: gb\_est27:\*  
105: gb\_est28:\*  
106: gb\_est29:\*  
107: gb\_est30:\*  
108: gb\_est31:\*  
109: gb\_est32:\*  
110: gb\_est41:\*  
111: gb\_est42:\*  
112: gb\_est43:\*  
113: gb\_est44:\*  
114: gb\_est45:\*  
115: gb\_est46:\*  
116: gb\_est47:\*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
```

```
em_gss_pln1:*
em_gss_pln2:*
em_gss_pro:*
em_gss_rod1:*
em_gss_rod2:*
em_gss_rod3:*
em_gss_rod4:*
em_gss_rod5:*
em_gss_vrt1:*
em_gss_vrt2:*
em_gss_vrt3:*
em_gss1:*
em_gss2:*
em_gss3:*
em_gss4:*
em_gss5:*
em_gss6:*
em_gss7:*
em_gss8:*
em_gss9:*
em_gss10:*
em_gss11:*
em_gss12:*
em_gss13:*
em_gss14:*
em_gss15:*
em_gss16:*
em_gss17:*
em_gss18:*
em_gss19:*
em_gss20:*
em_gss21:*
em_gss22:*
em_gss23:*
em_gss24:*
em_gss25:*
em_gss26:*
em_gss27:*
em_gss28:*
em_gss29:*
em_gss30:*
em_gss31:*
em_gss32:*
em_gss33:*
em_gss34:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	428.6	38.3	522	120	AW855262	AW855262	PM1-CT026
2	293.4	26.2	508	218	AZ283358	AZ283358	RPC1-23-1
C 3	200.2	17.9	461	15	AI045155	AI045155	UI-R-C1-k
4	168.6	15.1	312	12	AA790757	AA790757	vw18q09.r
5	134.4	12.0	935	106	AL531903	AL531903	AL531903
C 6	130.2	11.6	717	229	CNS01RJ1	CNS01RJ1	AL163990 Tetraodon
C 7	126.6	11.3	492	13	AA889777	AA889777	al50h01.s
8	120.2	10.7	485	115	AW447807	AW447807	BE751067 89973 MAR
9	112.6	10.1	549	139	BE751067	BE751067	BE751067 202785 MA
10	112.2	10.0	489	117	AW595507	AW595507	AW595507 fk31f08.y
11	108.4	9.7	522	143	BF073678	BF073678	BF073678 220845 MA
12	108.4	9.7	585	137	BE627479	BE627479	BE627479 uu52b07.y
13	106.2	9.5	577	118	AW632094	AW632094	AW632094 91641 MAR
14	105.4	9.4	464	118	AW632366	AW632366	AW632366 92219 MAR
15	101.2	9.0	958	174	BG174412	BG174412	BG174412 602334232
16	100.8	9.0	514	173	BG093377	BG093377	BG093377 ut70q04.y
17	97	8.7	558	139	BE750883	BE750883	BE750883 202546 MA
18	97	8.7	628	112	AW227957	AW227957	AW227957 up19e03.y

```
19 95.8 8.6 561 150 BF591285 7h44g03.x
20 95.2 8.5 671 172 BG025793 602274794
21 95.2 8.5 728 107 AU117058 AU117058
22 95.2 8.5 858 169 BF797340 602256969
23 90 733 104 AL959485 fd10g03.y
24 89.2 8.0 450 157 W04836 za8if05.r1
25 89.2 8.0 551 143 BF081124 233997 MA
26 87.2 7.8 787 106 AU080004 AU080004
27 86.8 7.8 886 144 BF101953 601752919
28 86.4 7.7 557 118 AW669715 113242 MA
29 85 519 143 BF045645
30 84.6 7.6 574 206 A0461727
31 84.4 7.5 598 120 AW822616 uq13b02.y
32 84 7.5 544 139 BE751588 203764 MA
33 83.6 7.5 665 141 BE901237 601676059
34 83.2 7.4 944 106 AL523112
35 82.8 7.4 487 137 BE656336 UI-M-BHO-
36 82.8 7.4 995 231 CNS04ROE
37 82.6 7.4 493 167 BE487374 176082 BA
38 82.6 7.4 955 174 BG173867 60233928
39 82.2 7.3 822 107 AU139387
40 81.6 7.3 499 114 AK389565
41 81.4 7.3 488 114 AW403941 UI-HF-BK0
42 80.8 7.2 417 205 AQ341822
43 80.8 7.2 899 150 BF581225
44 80.4 7.2 580 31 AV708554
45 80.4 7.2 937 231 CNS03GSG AV708554 AV708554
AL243385 Tetraodon
```

ALIGNMENTS

```
RESULT 1
AW855262 522 bp mRNA EST 19-MAY-2000
LOCUS PM1-CT0267-170300-006-g08 CT0267 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW855262
ACCESSION AW855262.1 GI:7950955
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 522)
Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-pm1-ct0267-170
300-006-g08&t3=2000-03-17&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 521.
Location/Qualifiers
1..522
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

FEATURES

source

```
/clone_lib="CT0267"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site.1: SmaI; Site.2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 123 a 133 c 123 g 143 t
ORIGIN
Query Match 38.3%; Score 428.6; DB 120; Length 522;
Best Local Similarity 97.5%; Pred No. 3.9e-115;
Matches 467; Conservative 0; Mismatches 9; Indels 3; Gaps 3;
Qy 314 GCAATGCCACGTGCAAGTTGCTAAAGGCATCATGCCATCAACTTTAACTGGCGATGC 373
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 CGGATGTCGGTGCATGTTGCTAAAGGCATCATG-CATCAACTTTAACTGGCGATGC 79
Qy 374 TGCTCCTGACCTTGATAGCATGGACCGGTACATGCCATTGTACAGGCGACTAATCAT 433
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 80 TGCTCCTGACCTTGATAGCATGGACCGGTACATGCCATTGTACAGGCGACTAATCAT 139
Qy 434 TCCGGCTCGGATCCAGACACTACCGCGCAGCAAAATCATCTGCCCTTG-TTGTGTGGGG 492
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 140 TCCGGCTCGGATCCAGACACTACCGCGCAGCAAAATCATCTGCCCTTGTTGTGTGGGG 199
Qy 493 CTGTGAGTCATCATCTCCAGCTCAACTTTTGTCTTCAACCAAAAATACACACCCCAAGGC 552
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 200 CTGTGAGTCATCATCTCCAGCTCAACTTTTGTCTTCAACCAAAAATACACACCCCAAGGC 259
Qy 553 AGCGATGCTGTGAACCCCAAGTACCAAACTGTCTCGGAG-CCCATCAGGTGGGAAGTGCT 611
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 260 AGCGATGCTGTGAACCCCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGGAAGTGCT 319
Qy 612 GATGTGGGGCTTGAGCTACTCTTTGGTTTCTTATCCCTTTGATGTTGATATATTGG 671
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 320 GATGTGGGGCTTGAGCTACTCTTTGGTTTCTTATCCCTTTGATGTTGATATATTGG 379
Qy 672 TTACAGTTTCATTGTCAAAACCTTGGTGAAGTCAAGTCTAGAATTTCTAAAAGGCACAAAGCCAT 731
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 380 TTACAGTACATGTCAAAACCTTGGTGAAGTCAAGTCTAGAATTTCTAAAAGGCACAAAGCCAT 439
Qy 732 CCGTGAATCATAGCTGTGGTCTTGTGTTTCTTGGCTTGTGAGATTCCTCATCAACATGG 790
|| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 440 CCGTGAATCATAGCTGTGGTCTTGTGTTTCTTGGCTTGTGAGATTCCTCATCAACATGG 498
RESULT 2
AZ283358 508 bp DNA GSS 27-JUL-2000
LOCUS RPCI-23-141E7.TV RPCI-23 Mus musculus genomic clone RPCI-23-141E7,
DEFINITION DNA sequence.
ACCESSION AZ283358
VERSION AZ283358.1 GI:9525144
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 508)
REFERENCE
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1993)
COMMENT Other_GSSs: RPCI-23-141E7.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
```



```
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia)
) With a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5',
TGTACCAATCTGAAGTGGGCGCGCGAATGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
BASE COUNT 70 a 80 c 78 g 84 t
ORIGIN

Query Match 15.1%; Score 168.6; DB 12; Length 312;
Best Local Similarity 71.4%; Pred. No. 1.2e-38;
Matches 222; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 424 ACTAAGTCATTCGGCTCCGATCCAGAACACTACCGCGAGCAAAATCATCTGCTTGT 483
DB 1 ACCAAGCTTTCCGGGTACGCTCCAGAACACTACGACAGTAAGCTCATCTGTATGGCA 60

QY 484 GTGTGGGGGCTGTAGTCATCATCTCCAGCTCAACTTTTGTCTTCAACCAAAATACAAC 543
DB 61 GTGTGGTTCATCTCCATCATCATCTCAAGCCCTACATTTAGCTTCAACAGAGATACGAG 120

QY 544 ACCCAAGCGAGCGATGCTGTGAAGCCCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGG 603
DB 121 CTGCAGGATCGTATGCTGTGAGCCCGGTACAGCTGTCTCAGAGCCCATCAGCTGG 180

QY 604 AAGTCGTGATGTTGGGGCTTGAGTACTCTTTGGTTTCTTTATCCCTTTGATGTTTATG 663
DB 181 AGGCTGCTGGGTATGGGACTGGAGCTGTTCTTTGGGTTCTTCCACCCCTTTGCTGTTATG 240

QY 664 ATATTTTGTACACGCTTCATTTGCAAAACCTTGGTGCAAGCTCAGAAATCTTAAAGGCAC 723
DB 241 GTGTCTGCTATCTGTTCAATAATCAGAACTTGGTGCGAGGCCGAGAACTTCCAGAGGCAC 300

QY 724 AAAGCCATCCG 734
DB 301 AGAACCATCCG 311

RESULT 5
AL531903 935 bp mRNA EST 13-FEB-2001
LOCUS AL531903 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM003YH06 5
DEFINITION prime, mRNA sequence.
ACCESSION AL531903
VERSION AL531903.1 GI:12795396
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 935)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. .935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DM003YH06"
/clone_lib="LTI_NFL001_NBC4"
```

```
Query Match 17.9%; Score 200.2; DB 15; Length 461;
Best Local Similarity 74.1%; Pred. No. 6.9e-48;
Matches 266; Conservative 0; Mismatches 90; Indels 3; Gaps 1;

QY 745 GCTGTGCTGCTGTGTTCTGCTGTGTCAGATTCCTCAATCAATGTTGCTTGTGACG 804
DB 460 GCTGTGCTGCTGTTCTGCTGCTGTCAGATTCCTCAATCAATGTTGCTTGTGACT 401

QY 805 GCTGTCAATTTGGTAAATGACCAATCTGTCAGAGCGAAAGCTAATGCTATACG 864
DB 400 GCAGCAACACAGCGCAAAATGGCGAGCTGCGAGCGCCGAGAAAGCCCTCGCCTACGCC 341

QY 865 AAAACTGTACAGAACTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
DB 340 AGAATGTGCTGAGTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 281

QY 925 TTTATTGGGCAAGTTCAGAACTACTTTCTGAAGATCTTGAAGACCTGTGCTGTGTG 984
DB 280 TTTATTGGGCAAGTTCAGAACTACTTTCTGAAGATCTTGAAGACCTGTGCTGTGTG 221

QY 985 AGAAGGAAGTACAGTCTCAGAGCTTCTCTGTCGCGGGAGGTAATCAGAAA---ACATT 1041
DB 220 AGGAGGAAGAGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 161

QY 1042 TCTCGCAGCAGCAGTCCAGCCGAGATACGACAACTGCTGCTGCTGCTGCTGCTGCTG 1100
DB 160 TCCAGGACAGCAGTGGAGACTGTAGAAATGACAACGATCGCTGCTTACCATGTAAAC 102

RESULT 4
AA790757 312 bp mRNA EST 06-FEB-1998
LOCUS AA790757 Soares_mammary_gland_NDMWG Mus musculus cDNA clone
DEFINITION vwi8909.r1 Soares_mammary_gland_NDMWG Mus musculus cDNA clone
IMAGE:1244224 5' similar to SW:CKR6_HUMAN P51684 C-C CHEMOKINE
RECEPTOR TYPE 6; , mRNA sequence.
ACCESSION AA790757
VERSION AA790757.1 GI:2850877
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 312)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-RHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-RHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:657912
Possible reversed clone: similarity on wrong strand
Seq primer: -28m3 rev2 Et from Amersham
High quality sequence stop: 217.
Location/Qualifiers
1. .312
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1244224"
/clone_lib="Soares_mammary_gland_NDMWG"
/sex="male"
```



```

QY 490 GGGCTGTGAGTCATCTCCAGCTCAACTTTTGTCTTTCAACCAAAAAATACAACACCCAA 549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 GTGGCGCGCTCATCTCTCGATCGCGAGATGAGGTACACCGGTCAACAACACACAGC 218
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 550 GGCAGCATGTCTGTGAACCAAGTACAAACTGTCTCGGAGCCCATCAGGTGGAAGCTG 609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 TGCAGCGCTTACACCGCAGCAAGACAGTCGCGCTC----- 178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 610 CTGATGTTGGGGCTTGAGCTACTCTTGGTTCTTTATCCCTTTTCATGCTTCATGATATT 569
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 ATCATCCAGGTGGCCAGATCGTCTGCCCTTCGCCCTCCCGCTGTAATCATGAGCATC 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 670 TGTACAGCTTCAATGTCACAAACCTTGGTCAAGCTCAGAACTTCATAAAGGCACAAAGCC 729
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 TGCTACAGCAGCATCATCAAAACCTCTGTCAAGCTCAGAACTTGTGCGGNAACAAGGCC 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 730 ATCCGTGTAATCATGCTGGTGGCTGTGTGTTCTGGCTTTCGATTCCT 780
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 ATCAAGTCACTCTGGCGGTGGTGGCGGTTTCTCTGCTGAGCCAGTTCCT 9
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AA889777/c
LOCUS
DEFINITION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
  1..492
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:1460785"
  /lab_host="Soares_NFL_T_GBC_S1"
  /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19W, testis NHT, and B-cell
NCL1CGAP.GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo.
134 a 119 c 117 g 122 t

Query Match 11.3%; Score 126.6; DB 13; Length 492;
Best Local Similarity 55.4%; Pred. No. 2.6e-24;
Matches 255; Conservative 0; Mismatches 199; Indels 6; Gaps 1;

```

```

Best Local Similarity 80.8%; Pred. No. 3.4e-26;
Matches 172; Conservative 0; Mismatches 39; Indels 2; Gaps 2;

QY 907 AACCCCTGTCTACGCTTTTATTGGCGAGAAGTTTCAGAACTACTTCTGAGATCTTG 966
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 AACCCCTGTCTTAAACCCCTTTTATTGGCCAGAAGTTCCGAAACTACTTTCGGAAGATCTTG 433
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 967 AAGGACCTGTGTGTGAGAAAGAGTACAAGTCTCAGGCTTCCTGTCGCGGAGG 1026
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 AA-GACCCCTGTGTCTCGAAGGAAGTACCAGTCTCAGGCTTCCCTTTCCCGGAGT 374
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1027 TACTCAGAAACATTTCTCGGAGACAGTACAGCCGAGATACACGACATGCGTCTGTC 1086
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 TCCCGGAAACCTTTTCTCGGAAGACCATTTAGACCGCAGTTACCGACATGGCGCGCTCC 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1087 TTCACCTATGTATAGAAAGCTGAGTCTCCCTAA 1119
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 -TCACCATTTGATAGAAGGTGGGCCCCCTAA 282
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
AW447807
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
  1..485
  /organism="Bos taurus"
  /db_xref="taxon:9913"
  /clone_lib="MARC 1BOV"
  /tissue_type="pooled"
  /lab_host="DH10B"
  /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
BASE COUNT 88 a 160 c 128 g 108 t
ORIGIN

Query Match 10.7%; Score 120.2; DB 115; Length 485;
Best Local Similarity 55.4%; Pred. No. 2.6e-24;
Matches 255; Conservative 0; Mismatches 199; Indels 6; Gaps 1;

```

```

QY 340 GGCATCTATGCCATCACTTAACCTGCGGAGTCTCTCTGACTTCATTAGATGAC 399
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 6 GCATCTACAGATAGCTTCTTCAGTGCATGCTCCTGCTGCTATGATCAGCATCGAC 65
QY 400 CGGTATACGCGCATGTATACAGCGACTAAGTCATTCGGCTCCGATCCAGAACACTACCG 459
Db 66 CGCTAGTTGGCATCGTCCAGCGCGTCTCGGCCACCGCCACCGTGCCTCTTC 125
QY 460 CCAGCAAAATCATCTGCTTGTGTGGGGCTGTGAGTCATCATCTCCAGCTCAACT 519
Db 126 ATCAGCAAGCTCTCTGCTGGGCATCTGGATGCTGCCATATGCTCTCCACCCAGAG 185
QY 520 TTGTGCTTCAACCAAAATACACACCCCAAGCAGCGATGCTGTGAACCCACAGTACCAA 579
Db 186 GTGATGACAGCGGATCCAGAAGA-----GCAGCAGCAGCAGCAGTGGGGTCTCC 239
QY 580 ACTGCTCTCGAGCCCATCAGGTGGAAGCTGCTCATGTTGGGGCTTGAGCTACTCTTTGGT 639
Db 240 CTTGTACCGACGACGTGGAGCCCTTGATCACCATCCAGTGGCCAGATGGTGTAGGC 299
QY 640 TTCCTTATCCCTTTGATGTTATGATATATTTGTTACAGCTTATTTGTTCAAAACCTTGGTG 699
Db 300 TTCCTGATCCCTCTGATGCGCATGAGCTTCTGTACTCTGTCATCATCCGACCCCTGCTC 359
QY 700 CAAGCTCAGAAATCTAAAGGACCAAGCCATCCGTTAATCATAGCTGTGGTGTGTG 759
Db 360 CAGGACGCAACTTCGAGCGCAACAGGCCATCAAGGTATCATATTGCTGTGCTGGTGC 419
QY 760 TTCTTGCTGTGTCAGATTCCTCATACATGGTCTCTCTTG 799
Db 420 TTCGTAGCCTTCCAGCTGCCTTAAGCGNGGTGTTCTG 459

RESULT 9
BE751067 549 bp mRNA EST 15-SEP-2000
LOCUS BE751067 202785 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BE751067
ACCESSION BE751067
VERSION BE751067.1 GI:10165059
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 549)
AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
,W.W. and Keele,J.W.
Design and use of four pooled tissue normalized cDNA libraries for
EST discovery in cattle
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smiththemal.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAAT
BACKWARD: GTTTCACGTCACGACG
Plate: 42 row: K column: 15
Seq primer: ATTTAGTGACACTATAG.
FEATURES
source
1..549
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/Note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, thymus,

```

```

semitendonosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
BASE COUNT 93 a 182 c 140 g 134 t
ORIGIN
Query Match 10.1%; Score 112.6; DB 139; Length 549;
Best Local Similarity 56.9%; Pred. No. 4.7e-22;
Matches 227; Conservative 0; Mismatches 169; Indels 3; Gaps 1;
QY 601 TGAAGCTGCTGATGTTGGGCTTGAGCTACTCTTTTGGTTTCTTATCCCTTTGATGTT 660
Db 115 TGGGAATGGTATGAGGGCTCTCCCCAGACCTTTTGGCTTCTCTCCCTCGCTGGTTC 174
QY 661 ATGATATTTTGTACAGCTTCATTGTCAAAACCTTGGTGCAGGCTCAGAAATCTAAAAG 720
Db 175 ATGCTGTTCTGCTACGATTCACCTCGCACGCTGTTTTCAGCCCAATGGGCGAGAAG 234
QY 721 CACAAAGCCATCCGTGTAAATCATAGCTGTGCTTGTGTTTCTGGCTGTGCAGATTCT 780
Db 235 CACGGGCGCATGCGGGTCATCTTTTGTCTGCTGCTCTTCTCTCTGCTGCTGGCTGCC 294
QY 781 CATAACATGTCCTGCTTGTGACGGCTGCTAATTTGGTAAATGA---ACGATCCTGC 837
Db 295 TACAACCTGTCTGATGTCGTCGACACCTCATGAGGGCCCATGATGTCGAGACCTGT 354
QY 838 CAGAGCGAAAAGCTTAATTTGGCTATACGAAACTGTACAGAACTCTCGCTTTCCTGCAC 897
Db 355 CAGCGCGCAACGACATTTGCGGGCGCTGATGCCACCGAGATCCTGGCTTCTCTGCAC 414
QY 898 TGTGCTGTAACCTGTGCTCTAGCTTTTATTGGGAGAAGTTTCAGAAACTTCTTCTG 957
Db 415 AGCTGCTCAACCCCTCATCTAGCTTTCATTGGCCAGAAGTTTCGCCACGAGTCTCT 474
QY 958 AAGATCTTGAAGGACCTGTGTGTGAGAGGAAGTAC 996
Db 475 AAGATCATGGCCATCCATGGCTGATCAGCAAGAGTTC 513

RESULT 10
AW595507 489 bp mRNA EST 22-MAR-2000
LOCUS fk31f08.y1 zebrafish fin day1 regeneration Danio rerio cDNA 5',
DEFINITION similar to SW:CKR6_HUMAN P51684 C-C CHEMOKINE RECEPTOR TYPE 6 ;
ACCESSION AW595507
VERSION AW595507.1 GI:7282765
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Rasbora; Danio.
REFERENCE 1 (bases 1 to 489)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Other_ESTs: fk31f08.x1
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@wustl.edu
cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:

```

infoGenomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de)

Seq primer: T3 ET from Amersham High quality sequence stop: 375.

FEATURES

Location/Qualifiers

1. .489

/organism="Danio rerio"

/db\_xref="taxon:7955"

/clone\_lib="zebrafish fin day1 regeneration"

/sex="mixed male and female"

/tissue\_type="1 day fin regenerates"

/lab\_host="E. coli XL0LR"

/note="Vector: pBK-CMV; Site:1: EcoRI; Site:2: XhoI; 1st strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed by second strand synthesis, and ligated to 5' adapter (5'-aattcgacgag-3', 3'-gccgtgctc-5'. cDNA was cloned directionally (EcoRI/XhoI) into Stratagene zap express lambda phage arms. Mass invivo excision done to obtain inserts in pBK-CMV phagemid."

BASE COUNT 98 a 125 c 132 g 133 t 1 others

Query Match 10.0%; Score 112.2; DB 117; Length 489;

Best Local Similarity 57.2%; Pred. No. 5.9e-22;

Matches 223; Conservative 0; Mismatches 164; Indels 3; Gaps 1;

QY 626 AGTACTCTTTGGTTCTTTATCCCTTTGATGTCATGATATTTTGTACACGTTTCATTG 685

DB 14 AGATGGTGTGGTTCTGTGCGCGTGTATGTTGTTCTGTCTACTCATGGACCA 73

QY 686 TCAAAACCTTGGTGCAGCTCAAAATCTAAAGGACACAAAGCCATCCGTGTAATCATAG 745

DB 74 TGTGACTCTTCTTAAGGGCGACAACTTTTCAGAGCACAAAGCGGTGCGTGTGTCCTGG 133

QY 746 CTGTGGTCTGTGTCTGTGCTGTGCAGATTCCTCATACATGTTCTCTCTGTGTGACGG 805

DB 134 CTGTGTTTTCGTGTGTGCTGTGCGACCTGCGGTATACAGTAGCATCTCGTGTGTACA 193

QY 806 CTGCTAATTTGGTAAATGAACCGATCTCTGCCAGAGCGAAAGCTAATTTGGCTATACGA 865

DB 194 CCAGTAACTGTTT---GTGAGAGAGTGTGCGGGAGGAGGAGGTCACGCTTATGACGC 250

QY 866 AACTGTACAGAAGTCTGCTGTTTCTGACATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 925

DB 251 TTTTCGGTTAGCCGACCGACGCGTATCTGCACTGCTGCTCAATCCATCCCTCTACGCT 310

QY 926 TTATTTGGCAGAGTTTCAGAACTACTTCTGAAGATCTTGAAGGACCTGTGTGTGTGA 985

DB 311 TCATCGGGGTCAAATTCAGAAGCCATCTGCGCAGATCTGCGAGATCTGCGGTGCTTC 370

QY 986 GAAGGAAGTACAGTCTCTCAGGCTTCTCCT 1015

DB 371 GGAGAGTACATCTACTCANGACGCTCCT 400

RESULT 11

BF073678

LOCUS

BF073678

DEFINITION

220645 MARC 280V Bos taurus cDNA 5', mRNA sequence.

ACCESSION

BF073678

VERSION

BF073678.1

KEYWORDS

EST.

SOURCE

COW.

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 522)

REFERENCE

Smith,T.P.L., Casas E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred

v0.980904.e. Vector identified by cross\_match with the -minscore 18

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCAGTCAGCAGC

Plate: 83 row: H column: 13

Seq primer: ATTTAGGTGACACATATAG.

Location/Qualifiers

1. .522

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 280V"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SPORT6; Site:1: XbaI; Site:2: XhoI;

Library made from pooled tissue from testis, thymus,

semimendonsus muscle, longissimus muscle, pancreas,

adrenal, and endometrium."

BASE COUNT 97 a 182 c 137 g 106 t

ORIGIN

Query Match 9.7%; Score 108.4; DB 143; Length 522;

Best Local Similarity 55.1%; Pred. No. 7.9e-21;

Matches 236; Conservative 0; Mismatches 186; Indels 6; Gaps 1;

QY 372 GCTGCTCTGACTTCATTTAGCATGGACGGTACATCGCCATTGTACAGCGCAGTAAGTC 431

DB 1 GCTCCTCTGTATGATCAGCATCGCCGCTACGTTGCCATCGTCCAGCCGCTCTCGGC 60

QY 432 ATTCGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTTGTGTGGGG 491

DB 61 CCACGCCACCGTGCCGCTCTCTCATCAGCAGCTCTCTGCTGGCATCTGGAT 120

QY 492 GCTGTGATCATATCTCCAGCTCACTTTTGTCTTCAACCAAAATACACACCAAGG 551

DB 121 GCTGCCATAGTGTCTCCACCCAGAGGTGATGTACAGCGGATCCAGAAGA-----G 174

QY 552 CAGGATGCTGTGAACCCAGTACCAAACTGCTCGGAGCCCATCAGTGGAGCTGCT 611

DB 175 CAGCAGCAGCAGGCACTCGCGTCTCTCGTCCAGCAGCAGTGGAGCCCTTGATCAC 234

QY 612 GATGTTGGGCTTGAGCTACTCTTTGGTTTCTTTATCCCTTTGATGTTTCATGATATTTG 671

DB 235 CATCCAGTGGCCAGATGGTGGTAGGCTTCTGATCCCTCTGATGCTGATGCTGCTG 294

QY 672 TTACACGTTTCATTTGCAAAACCTTGGTGAAGCTCAGAAATTTCTAAAGGCACAAAGCCAT 731

DB 295 CTACCTCGTCATCATCCGACCCCTGCTCCAGCAGCAGCACTTCGAGCGCAACAGGCCAT 354

QY 732 CCGTGAATCATAGTGTGGGCTGTGTTTCTGGCTTCTGAGATTCCTCATTAACATGGT 791

DB 355 CAAGGTGATCATTTGTTGGTGTGCTCTCTGCTAGCTTCCAGCTCCCTACAAACGGGT 414

QY 792 CCTGCTT 799

DB 415 GGTTCGTG 422

RESULT 12

BE627479

LOCUS

BE627479

585 bp

mRNA

EST

24-AUG-2000

[illegible]

Qy 131 ACTCCTTGATCTGTGTTCTTGGCCCTCCTGGGAAATATCTGGTGGTGATCACCTTTGCTT 190

Db 193 ACTCATCAATTTGCTTCGTGGGCCCTTCCTGGGCAACGGGCTGGTCATCTGACCTACATCT 252

BASE COUNT	227 a	266 c	266 g	199 t	Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH
227 a	266 c	266 g	199 t		

QY 119 ACCGAATGGCTACTCCTTGATCTGTCTCTGGCCCTCTGGGGAATATTTCTGGTGTTGA 178

```
Db 195 GCC-CACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCC 253
QY 179 TCACCTTTTGCTTTTATAGAGGCCAGGCTCATGACAGACGCTCTATCTCTTTGAACATGG 238
Db 254 TGGTCATGGGTTACCAGAGAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGT 313
QY 239 CCATTGCAGACATCCTCTTTGTCTTACTCTCCCATTTCTGGGCACTGAGTCATGCCACTG 298
Db 314 CAGTGGCTGACCTCCTCTTTGTGCATCACACTCCCTTCTGGGCACT--TGATGCCATGG 370
QY 299 GTGCGTGGGTTTTCAGCAATGCCAGTGCAAGTTGCTAAAGGCATCTATGCCATCAACT 358
Db 371 CTGACTGGTACTTTGGGAA-ATTTTGTGAAGGCTGTCCATATCATCTACACTGTCAACC 429
QY 359 TTAACGTGGGGATGCTGCTCCTGACTTGCAATTAGCATGGACCGGTACATCGCCATTGTAC 418
Db 430 TCTACAGCAGCGTTCTCATCCTCGGCTTCATCAGCTGGACCGGTACCTCGCCATTGTCC 489
QY 419 AGGCGACTAA 428
Db 490 ACGCCACCAA 499
```

Search completed: May 24, 2001, 07:59:00  
Job time: 4086 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2001, 06:54:34 ; Search time 77.37 Seconds  
(without alignments)  
2525.273 Million cell updates/sec

Title: US-08-887-977-9  
Perfect score: 1119  
Sequence: 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 302621 seqs, 87301344 residues

Total number of hits satisfying chosen parameters: 605242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	222	19.8	1900	1 US-08-153-848-18	Sequence 18, Appl
2	222	19.8	1900	4 US-09-299-843A-18	Sequence 18, Appl
3	222	19.8	1900	5 PCT-US93-11153-18	Sequence 18, Appl
4	222	19.8	2058	1 US-08-153-848-6	Sequence 6, Appl
5	222	19.8	2058	4 US-09-299-843A-6	Sequence 6, Appl
6	222	19.8	2058	5 PCT-US93-11153-6	Sequence 6, Appl
7	222	19.8	2160	1 US-08-153-848-14	Sequence 14, Appl
8	222	19.8	2160	4 US-09-299-843A-14	Sequence 14, Appl
9	222	19.8	2160	5 PCT-US93-11153-14	Sequence 14, Appl
10	216.2	19.3	2085	4 US-09-299-843A-65	Sequence 65, Appl
11	215.6	19.3	2154	1 US-08-383-750-1	Sequence 1, Appl
12	215.6	19.3	2154	3 US-08-383-751A-1	Sequence 1, Appl
13	215.6	19.3	2154	3 US-08-352-678-1	Sequence 1, Appl
14	215.6	19.3	2154	5 PCT-US93-09636-1	Sequence 1, Appl
15	206.4	18.4	2751	1 US-08-153-848-23	Sequence 23, Appl
16	206.4	18.4	2751	4 US-09-299-843A-23	Sequence 23, Appl
17	206.4	18.4	2751	5 PCT-US93-11153-23	Sequence 23, Appl
18	181.4	16.2	2156	1 US-08-012-988A-1	Sequence 1, Appl
19	177.6	15.9	1200	5 PCT-US95-03032-1	Sequence 1, Appl
20	174.2	15.6	1607	4 US-08-875-573-19	Sequence 19, Appl
21	173.2	15.5	1373	5 PCT-US92-02977-6	Sequence 6, Appl
22	173.2	15.5	1373	5 PCT-US95-03032-3	Sequence 3, Appl
23	172.8	15.4	1200	5 PCT-US92-02977-1	Sequence 1, Appl
24	171.2	15.3	1510	1 US-07-759-568-4	Sequence 4, Appl
25	171.2	15.3	1748	1 US-08-202-056-8	Sequence 8, Appl
26	166	14.8	1106	5 PCT-US92-02977-5	Sequence 5, Appl
27	166	14.8	1106	5 PCT-US95-03032-4	Sequence 4, Appl

28	163.4	14.6	1176	5 PCT-US95-03032-2	Sequence 2, Appl
29	163.4	14.6	1883	1 US-08-202-056-2	Sequence 2, Appl
30	163.4	14.6	1933	1 US-08-076-093A-1	Sequence 1, Appl
31	163.4	14.6	1933	1 US-08-410-451-1	Sequence 1, Appl
32	163.4	14.6	1933	1 US-08-410-455-1	Sequence 1, Appl
33	163.4	14.6	1933	1 US-08-418-919-1	Sequence 1, Appl
34	163.4	14.6	1933	1 US-08-410-453A-2	Sequence 2, Appl
35	163.4	14.6	1933	1 US-08-701-265-1	Sequence 1, Appl
36	163.4	14.6	1933	1 US-08-410-454A-2	Sequence 2, Appl
37	163.4	14.6	1933	2 US-08-284-586-1	Sequence 1, Appl
38	163.4	14.6	1933	2 US-08-410-456A-2	Sequence 2, Appl
39	163.4	14.6	1933	2 US-08-805-478-1	Sequence 1, Appl
40	163.4	14.6	1933	2 US-08-802-627A-1	Sequence 1, Appl
41	163.4	14.6	1933	2 US-08-801-238-1	Sequence 1, Appl
42	163.4	14.6	1933	2 US-08-801-238-1	Sequence 1, Appl
43	163.4	14.6	1933	3 US-09-104-296-1	Sequence 1, Appl
44	163.4	14.6	1933	5 PCT-US94-06380-1	Sequence 1, Appl
45	154.8	13.8	1094	3 US-09-178-637-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-08-153-848-18  
; Sequence 18, Application US/08153848  
; Patent No. 5759804  
; GENERAL INFORMATION:  
; APPLICANT: Godiska, Ronald  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Schweikart, Vicki L.  
; TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,848  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,452  
; FILING DATE: 17-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5759804and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31794  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1900 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: 1..168  
; FEATURE:  
; NAME/KEY: exon

LOCATION: 169..1245  
FEATURE: CDS  
NAME/KEY: 169..1242  
LOCATION: 169..1242  
FEATURE: 3'UTR  
NAME/KEY: 1246..1900  
LOCATION: 1246..1900  
US-08-153-848-18

Query Match 19.8%; Score 222; DB 1; Length 1900;  
Best Local Similarity 55.5%; Pred. No. 7.3e-56;  
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

QY 71 GTTACTGTGCTCTTCCAGAGGTGAGGAGTCTCCAGGCTATTGTACCGAATTCCT 130  
DB 243 GTCTTTGTGCTCAAGAAGAGGAGTGGGAATTT-TAAAGCTTGGTTCCTTATCATGT 301  
QY 131 ACTCCTGTATCTGTTCTTGGCCCTCCGGGAATATCTGTGTGATCAGCTTTGCTT 190  
DB 302 ACTCATCATATTTGTTGCGGCGCTTACGGCAATGGGCTGTGTTGAGCTATATCT 361  
QY 191 TTTATAAGAAGCCAGGTCTATGACAGCTCTATCTCTTTGAACATGGCCATTGCAGACA 250  
DB 362 ATTCAAGAGGCTCAAGACCATGACGATACCTACTGCTCAACCTGGCGGTGGCAGACA 421  
QY 251 TCCTCTTTGTTTACTCTCCATTTCTGGGAGTGGAGTATGCCACTGTGGTGGGTTT 310  
DB 422 TCCTCTCTCTGCTGACCTTCTGCTTCTGGGCTACAGCGCGCCA---AGTCTGGGTCT 478  
QY 311 TCAGCAATGCCAGTGCAGTTGCTTAAAGGATCTATGCCATCACTTAACTCGGGA 370  
DB 479 TCGGTGTCACATTTTGCAGAGCTCATTTTGGCCATCTACAAGATGAGCTTCTTCAGTGGA 538  
QY 371 TCTGCTCTGCTGACTTGTAGTACGAGCGGTGATCATCGCCATTGTACAGGCGACTAAGT 430  
DB 539 TCGTCTACTTCTTGTGATCAGCATGACGCTACGTGCGCATGCTCCAGGCTGTCTCAG 598  
QY 431 CATTCGGGTCGATCCAGAACTACGCGGAGCAAAATCATCTGCTTGTGTGGG 490  
DB 599 CTCACGGGCGGCGGCTCTCTCATCAGCAAGCTGTCTGTGGGCTATCGGA 658  
QY 491 GGCTGTGATCATCTCCAGCTCACTTTTGTCTTCAACCAAAATACACACCCCAAG 550  
DB 659 TACTAGCCACAGTGTCTCTCAACAGAGCTCTCTACAGTACCTCCAGAGA----- 712  
QY 551 GCAGCGATGCTGTGAACCCAAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGAAGCTGC 610  
DB 713 GCAGAGTGAAGCAGGATGCGATGCTCTCATCAGAGAGCATGTGGAGGCTTTATCA 772  
QY 611 TGATGTGGGCTTGAAGTACTTGTGTTTCTTTATCCCTTTGATGTTTATGATTTT 670  
DB 773 CCATCCAGTGGCCCGAGATGTTGATCGGCTTCTGTGCTCCCTGTGGCATGAGCTTCT 832  
QY 671 GTTACAGTTTCAATGTCAAAACCTTGGTCAAGCTCAGATTTCTAAAGGCACAAAGCCA 730  
DB 833 GTTACCTTGTATCATCTCCAGCCCTGCTCCAGGCGACGCAACTTGGAGCGCAACAAGGCCA 892  
QY 731 TCGGTGTAATCATAGTGTGTGCTTGTGTTTCTTGTGCTTGTGATTTCTCATAACA--- 787  
DB 893 TCAAGGTGATCATGCTGTGGTGTGGTCTTCAATAGTCTTCCAGCTGCCCTCAATGGGG 952  
QY 788 TGGTCTGCTTGTGAGGCTGCTTAATTTGGGTAAATGAACCGATCCTGCCAGAGCGAAA 847  
DB 953 TGGTCTGCGCCAGAGGTGGCCACTTCAACATCACCAGTAGCAGCTGTGAGCTCAGTA 1012  
QY 848 AGCTAATTTGGCTTATACGAAACTGTACAGAAAGTCTGCTGCTTCTGCTGCTGCTGA 907  
DB 1013 AGCAACTCAACATCGCTACGACGTCACCTACAGCTGCGCTGCTGCTGCTGCTGCTCA 1072  
QY 908 ACCCTGTGCTTACGCTTTTATTTGGGAGAGTTCAGAAACTACTTCTGAAGATCTTGA 967  
DB 1073 ACCCTTCTGTGACGCTTCTATCTCCGCGGTCAAGTTCGCAACGATCTCTTCAAGCTTCA 1132

QY 968 AGGACCTGTGTGTGTGAGAGGAAGTA 995  
DB 1133 AGGACCTGGGCTGCTCAGCAGGAGCA 1160  
RESULT 2  
US-09-299-843A-18  
Sequence 18, Application US/09299843A  
Patent No. 6107475  
GENERAL INFORMATION:  
APPLICANT: Godiska, Ronald  
APPLICANT: Gray, Patrick W.  
APPLICANT: Schweikart, Vicki L.  
TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/299,843A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/088,337  
FILING DATE: 01-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/153,848  
FILING DATE: 17-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,452  
FILING DATE: 17-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Jill E. Uhl  
REGISTRATION NUMBER: 43,213  
REFERENCE/DOCKET NUMBER: 27866/32059B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1900 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1..168  
FEATURE:  
NAME/KEY: exon  
LOCATION: 169..1245  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 169..1242  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 1246..1900  
US-09-299-843A-18  
Query Match 19.8%; Score 222; DB 4; Length 1900;

Best Local Similarity 55.5%; Pred. No. 7.3e-56;  
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

Qy 71 GTTACTGTGCTCTTGCAGGAGGTCTTCCAGGCTATTTGTACCGAATTGCCT 130  
Db 243 GTCCTTGTGCTCCAGAGGAGCGTGGGAACCTT-TAAGCCCTGGTTCCTCCCTATCATGT 301  
Qy 131 ACTCCTTGATCTGTGTTCTTGGCCCTCTGGGGAATATTTCTGGTGGTATCACCTTTGCTT 190  
Db 302 ACTCCATCATTTGTTCTGGGCTACTGGCAATGGCTGGTGTGTGACCTATATCT 361  
Qy 191 TTTATAAGAGGCGAGGTCTATGACAGAGCTCTATCTTGTACATGGCCATTTGCAGACA 250  
Db 362 ATTTCAAGAGGCTCAAGACCATGACCGGATACCTGCTCAACCTGGCGGTGGCAGACA 421  
Qy 251 TCCCTTTGTTCTTACTCTCCCATCTTGGCAGTGAGTATGACCTGCTGGTGGGTTT 310  
Db 422 TCCCTTTGTTCTTACTCTCCCATCTTGGCAGTGAGTATGACCTGCTGGTGGGTTT 310  
Qy 311 TCAGCAATGCCAGTGTCAAGTTGTAAAGGCTATCTATGCCATCAACTTTAACTGCGGGA 370  
Db 479 TCGGTGTCACCTTTTGAAGCTATCTTTGCCATCTCAAGATGAGCTTCTTCACTGGCA 538  
Qy 371 TGCTGCTCTGACTGTGATAGCAGGAGTGTATGATGCTGATGCTGATGCTGATGCTGATG 430  
Db 539 TGCTGCTCTGACTGTGATAGCAGGAGTGTATGATGCTGATGCTGATGCTGATGCTGATG 490  
Qy 431 CATTCGCGCTCCGATCCAGACACTACCGGCGCAGCAAAATCATCTGCTTGTGTTGGG 490  
Db 599 CTCACGCGCCCGGCTGCGGCTGCTTCTCATCAAGAGCTGTCTGTGGGCTATCTGGA 658  
Qy 491 GCCTGTGAGTCTATCTGCTGAGTCAACTTTTGTCTCAACCAAAATATAACACCCCAAG 550  
Db 659 TACTAGCCAGAGTCTCTCCATCCAGAGTCTCTGATGAGTCTCCAGGGA----- 712  
Qy 551 GCAGCGATGCTGTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGAAGTGC 610  
Db 713 GCAGCGATGAGCAAGCGATGCGATGCTCTCTCATCACAGAGCATGTGAGGCTTTATCA 772  
Qy 611 TGATGTTGGGCTGTGAGTCTACTCTTGGTGTCTTTATCCCTTTGATGTTCAATATTTT 670  
Db 773 CCATCCAGGTGGCCAGATGGTGTATCGGCTTCTGTGCTCCCTGCTGCTGCTGCTGCT 832  
Qy 671 GTTACGTTCTATGTTCAAAACCTTGTGCAAGCTCAGAACTTCTAAAGGCAACAAAGCCA 730  
Db 833 GTTACCTGTATCATCCGACCCCTGCTCCAGGCGCCCACTTTGAGCGCAACAAAGCCA 892  
Qy 731 TCCGTGTAATCATAGCTGTGCTGTGTTCTGTGCTGTTGTGCTGATTCCTCATAACA--- 787  
Db 893 TCAAGGTGATCATCGCTGTGCTGTGCTGCTTCTCATAGTCTTCCAGCTGCCCTACAAATGGG 952  
Qy 788 TGGTCTGCTGTGAGCGGCTGTAAATTTGGGTAAATGAACCGATCTGTCAGAGCGAAA 847  
Db 953 TGGTCTGCGCCCAAGGTGGCCAACTTCAACATCAGCATGAGCAGTGTGAGCTCAGTA 1012  
Qy 848 AGCTAATTTGCTATACGAAAGTGTACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 907  
Db 1013 AGCAACTCAACATCGCTAGAGCTACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1072  
Qy 908 ACCCTGTGCTCTACGCTTTTATTTGGGCGAGAGTTTCAGAAATCTTCTCAAGATCTTGA 967  
Db 1073 ACCCTTTTGTAGCGCTTCTATCGGCTGCAAGTTCAGCAAGTCTCTTCAAGCTCTTCA 1132  
Qy 968 AGGACCTGGGTGTGAGAGAGGAAGTA 995  
Db 1133 AGGACCTGGGCTGCTCAGCCAGGAGCA 1160

## RESULT 3

PCT-US93-11153-18

; Sequence 18, Application PC/TUS9311153

; GENERAL INFORMATION:

; APPLICANT: Godiska, Ronald

; APPLICANT: Gray, Patrick W.  
; APPLICANT: Schweikart, Vicki L.  
; TITLE OF INVENTION: Novel Seven Transmembrane Receptors  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/11153  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,452  
; FILING DATE: 17-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Noland, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31794  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1900 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: 1..168  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 169..1245  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 169..1242  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: 1246..1900  
; PCT-US93-11153-18

Query Match 19.8%; Score 222; DB 5; Length 1900;

Best Local Similarity 55.5%; Pred. No. 7.3e-56;

Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

Qy 71 GTTACTGTGCTCTTGCAGGAGGTCTTCCAGGCTATTTGTACCGAATTGCCT 130

Db 243 GTCCTTGTGCTCCAGAGGAGCGTGGGAACCTT-TAAGCCCTGGTTCCTCCCTATCATGT 301

Qy 131 ACTCCTTGATCTGTGTTCTTGGCCCTCTGGGGAATATTTCTGGTGGTATCACCTTTGCTT 190

Db 302 ACTCCATCATTTGTTCTGGGCTACTGGCAATGGCTGGTGTGTGACCTATATCT 361

Qy 191 TTTATAAGAGGCGGAGTCTATGACAGAGCTCTATCTTGTACATGGCCATTTGCAGACA 250

Db 362 ATTTCAAGAGGCTCAAGACCATGACCGGATACCTGCTCAACCTGGCGGTGGCAGACA 421

Qy 251 TCCCTTTGTTCTTACTCTCCCATCTTGGCAGTGAGTATGACCTGCTGGTGGGTTT 310

Db 422 TCCCTTTGTTCTTACTCTCCCATCTTGGCAGTGAGTATGACCTGCTGGTGGGTTT 310

; APPLICANT: Godiska, Ronald

QY 311 TCAGCAATGCCACGTGCAAGTTGCTAAAGGCACTATGTCACATCAACCTTTAACTGCGGGA 370  
Db 479 TCGGTGTCACCTTTTGGCAAGCTCATCTTTGGCCATCTACAAGATGAGCTTCTTCACTGGCA 538  
QY 371 TGCTGCTCCTGCTACTTATGATAGCAGCGGTATGATCGCCATTCGACAGGCACTAACT 430  
Db 539 TGTCTCTACTTCTTGGATCAGCATGACCGGTAGCTAGCTAGCGCCATCTCCAGGCTGCTCAG 598  
QY 431 CATTCGCGCTCCGATCCAGAACACTACCGCGAGCAAAATCATCTGCTTGTGTGGG 490  
Db 599 CTCACCGCACGTCGCCGCGCTCTCTCATCAGCAAGCTGCTCTGTGGGCACTCTGGA 658  
QY 491 GGCTGTAGTCATATCTCCAGCTCAACTTTTGTCTTCAACCAAAAATACACACCCCAAG 550  
Db 659 TACTAGCCACAGTGTCTCCATCCAGAGCTCTCTGACAGTACCTCCAGAGGA----- 712  
QY 551 GCAGGATGTCTGTCAACCCAGTAGTACCAAACTGCTCTCGAGGCCCATCAGGTGGAAGCTGC 610  
Db 713 GCAGGATGAGCAAGCATGCGATGCTCTCTCATCAGAGATGCTGGAGGCTTTATCA 772  
QY 611 TGATGTTGGGCTTGAGCTACTCTTTTGGTTCTTTTATCCCTTTGATGTTTATGATATTT 670  
Db 773 CCATCCAGGTGCGCCAGATGCTGATCGCTTCTGCTGCCCTGCTGGCCATGAGCTTCT 832  
QY 671 GTTACACGTTCAATGTCAAAACCTTGGTGCAAGCTCAGAAATCTAAAAGGCACAAAGCCA 730  
Db 833 GTTACCTTGTATCATCGCACCTGCTCCAGGCACGCAACTTTTGAGCGCAACAAAGGCCA 892  
QY 731 TCCGTGTAATCATAGCTGTGGTGTGTTTCTGCTGCTGTCAGATCTCTCATATAACA--- 787  
Db 893 TCAAGGTGATCATCTTGGTGTGGTCTTCTATAGTCTTCCAGCTGCTCCATCAATGGG 952  
QY 788 TGGTCTGCTGTGACGGCTGCTAATTTGGGTAAATGAACCGATCTCTGCCAGAGCGAAA 847  
Db 953 TGGTCTGCGCCAGACGGTGGCACTTCAACATCACCAGTAGCACCTGTGAGCTCAGTA 1012  
QY 848 AGCTAATTTGGTATACGAAACTGTACAGAAAGTCTTGGCTTCTGCTGCTGCTGCTGCTGA 907  
Db 1013 AGCAACTCAACATCGCTACGACGTCACTACAGCTGCGCTGCGCTGCTGCTGCTGCTGCTA 1072  
QY 908 ACCCTGCTCTACCTTTTATTGGGCAAGATTCAGAACTACTTCTGAGATCTTGA 967  
Db 1073 ACCCTTCTTGTACCTCTTCTATCGCGGTCAAGTTCGCAACGATCTCTTCAAGCTCTTCA 1132  
QY 968 AGGACCTGTGGTGTGAGAGGAAGTA 995  
Db 1133 AGGACCTGGCTGCTCAGCCAGGAGCA 1160

RESULT 4  
US-08-153-848-6  
; Sequence 6, Application US/08153848  
; Patent No. 5759804  
; GENERAL INFORMATION:  
; APPLICANT: Godiska, Ronald  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Schweikart, Vicki L.  
; TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,848  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,452  
; FILING DATE: 17-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5759804and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31794  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2058 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 166...1395  
; US-08-153-848-6

Query Match 19.8%; Score 222; DB 1; Length 2058;

Best Local Similarity 55.5%; Pred. No. 7.5e-56;

Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

QY 71 GTTACTGTCTCTTGCAGGAGGTTCAGGAGTCTTCTCCAGGCTATTTGTACCGAATTCGCT 130  
Db 396 GTCTTTGTCTCCAGAGGACGCTGCGGAACTT-TAAAGCCTGCTTCTCCCTATCATGT 454  
QY 131 ACTCTTGTATCTGTCTTGGCTCTCTGGGAATATTTCTGGTGGTATACACCTTTGCTT 190  
Db 455 ACTCCATCATTTTCTGCTGGGCTACTGGGCAATGGCTGGTGTGACCTATATCT 514  
QY 191 TTTATAAGAGGCGGCTCTATGACAGAGCTCTATCTTTGAACATGCCATTCAGACA 250  
Db 515 ATTCAAGAGGCTCAAGACCATGACCGATACCTACCTGCTCAACCTGGCGGTGCAGACA 574  
QY 251 TCCTCTTTGTTTACTCTCCCATTTCTGGCAGTGAATCATGCCACTGGTGGGTTT 310  
Db 575 TCCTCTCTCTGACCTTCCCTTCTGGGCTTACACGCGGCA---AGTCTGGGTCT 631  
QY 311 TCAGCAATGCCACGTGCAAGTGTCTTAAAGCATCTATGCCATCAACTTTAACTGCGGGA 370  
Db 632 TCGGTGTCCACTTTTGAAGCTCATCTTTGCCATCTACAAGATGAGCTTCTTCACTAGTGCA 691  
QY 371 TGCTGCTCTGACTTGCATTTAGCATGAGCGGTACATCGCCATTTGTACAGCGCACTAAGT 430  
Db 692 TGTCTCTACTTCTTGTGATCAGCATTTGACGCTACGTGGCCATGCTCCAGGCTGTCTCAG 751  
QY 431 CATTCGCGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTTGTGTGGG 490  
Db 752 CTCACGCCACCGTGGCCGCTCTTCTCATCAGCAAGCTGTCTGTGTGGGCACTGGA 811  
QY 491 GGCTGTGAGTCATCTCCAGCTCAACTTTTGTCTTCAACCAAAAATACAAACCCCAAG 550  
Db 812 TACTAGCCACAGTGTCTCCATCCAGAGCTCTGTACAGTACGCTCCAGAGGA----- 865  
QY 551 GCAGGATGTCTGTGAACCCCAAGTACCAACTGTCTCGGAGCCCATCAGGTGGAGCTGC 610  
Db 866 GCAGGATGTGAACCCCAAGTACCAACTGTCTCATCAGAGCATGTGGAGGCTTTATCA 925  
QY 611 TGATGTTGGGCTTGTGATCTCTTGGTCTTCTTATCCCTTTGATGTTTATGATATTTT 670  
Db 926 CCATCCAGGTGGCCAGATGGTGTCTTCTGGTGGCTGCTGGCCATGAGCTTCT 985  
QY 671 GTTACAGCTTCTATTGTCAAAACCTTGTGCAAGCTCAGAAATTTTAAAGGCACAAAGCCA 730

Db 986 GTTACCTGTGTCATCATCCGACCCCTGCTCCAGGCAGCAACTTTGAGCGCAACAAGGCCA 1045  
QY 731 TCCGTGTAATCATAGCTGTGGTCTTGTGTTTCTGGCTTGTGATTCCTCATACACA --- 787  
Db 1046 TCAAGTGATCATCGGTGTGTGCTGTCTCTATAGTCTTCCAGCTGCCCTACAAATGGG 1105  
QY 788 TGGTCTGCTTGTGAGCGCTGCTAAATTTGGGTAATAATGAACCGATCTGCCAGAGCGAAA 847  
Db 1106 TGGTCTGGCCAGAGCGGTGGCAACTTCACATCACCAGTACACCTGTGAGCTCAGTA 1165  
QY 848 AGCTAATGTGCTATACGAAACATGTCACAGAACTCTCTGGCTTCTTCTGCACTGCTGCCCTGA 907  
Db 1166 AGCAACTCAACATCGCTACGAGCTCACCTACAGCTGGCGCTGCTGCCGTCA 1225  
QY 908 ACCGTGCTCTAGCTTTTATTTGGCGAGAGTTTCAGAACTACTTTCTGAAGATCTTGA 967  
Db 1226 ACCCTTCTTGTAGCGCTTCATCGGCTCAAGTTCCGCAACGATCTCTTCAAGCTCTTCA 1285  
QY 968 AGGACCTGTGTGTGAGAGGAAGTA 995  
Db 1286 AGGACCTGGCTGCCCTCAGCCAGGACGA 1313

## RESULT 5

US-09-299-843A-6  
; Sequence 6, Application US/09299843A  
; Patent No. 6107475

## ; GENERAL INFORMATION:

; APPLICANT: Godiska, Ronald  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Schweikart, Vicki L.  
; TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,843A  
; FILING DATE:

## ; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 09/088,337

; FILING DATE: 01-JUN-1998

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/153,848

; FILING DATE: 17-NOV-1993

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/977,452

; FILING DATE: 17-NOV-1992

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Jill E. Uhl

; REGISTRATION NUMBER: 43,213

; REFERENCE/DOCKET NUMBER: 27866/32059B

## ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

## ; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2058 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 166..1395  
; US-09-299-843A-6

Query Match 19.8%; Score 222; DB 4; Length 2058;

Best Local Similarity 55.5%; Pred. No. 7.5e-56;  
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

QY 71 GTTACTGTGCTCCCTGTCAGGAGTTCAGGCAGTTCCTCCAGGCTATTGTACCGAAATGTCCT 130  
Db 396 GTCTTTGTGCTCCRAAGAAGAGCGTGGGAACCTT-TAAGCCCTGGTTCCTCCCTATCATCT 454  
QY 131 ACTCCTTGATCTGTGTTCTTGGCCCTCCCTGGGAAATATTCTGGTGTGATCATCTTTGCTT 190  
Db 455 ACTCCATCATTTGTTCTGTTGGGCTTACTTGGGCAATGGGCTGGTCTGTGTGACCTATATCT 514  
QY 191 TTTATAAGAGGCCAGGTCTATGACAGAGCTCTATCTTGAACATGGCCATTGCAGACA 250  
Db 515 ATTTCAAGAGGCTCAAGACCATGACCGATACCTGCTCAACCTGGCGGTGGCAGACA 574  
QY 251 TCCTCTTTGTTTACTCTCCCACTTCTGGGAGTGCAGTCATGCCACTGGTGGTGGGTTT 310  
Db 575 TCCTCTTCTCTGACCCCTTCCCTTCTGGGCTACAGCGCGGCCA---AGTCTGGGTCT 631  
QY 311 TCAGCAATGCCACGTGCAAGTTGCTTAAAGGCATCTATGCCATCACTTAACTTAACTGGGGA 370  
Db 632 TCGGTGTCCACTTTTGGCAAGCTCATCTTTGCCATCTACAAGATGAGCTTCTTTCAGTGGCA 691  
QY 371 TGTGCTCTCTGACTTGCATTAGCATGGACCGGTACATGCCATTCGCAATGTACAGGCGCACTA 430  
Db 692 TGCTCTACTTCTTTGTCATGACATGACCGTACCGCTACGTCGTCGTCGTCGTCGTCAG 751  
QY 431 CATTCCGGCTCCGATCCAGAACACATACCGCGCAGCAAAATCATCTGCTTGTGTTGTGGG 490  
Db 752 CTCACCGCCACCGTCCCGCGCTCTCTTCATCAGCAAGCTGCTCTGTGTGGGCACTGGA 811  
QY 491 GGCTGTGATCATCATCTCCAGCTCAACTTTGCTTCAACCAAAATAACAACCCAAAG 550  
Db 812 TACTAGCCACAGTGTCTCCATCCACAGAGTCTCTACAGTCACCTCCAGAGA----- 865  
QY 551 GCAGGATGTCTGTGAACCCCAAGTACCAAACTGTCTGGAGGCCATCAGTGGAGCTGC 610  
Db 866 GCAGCAGTGAGCAAGCGATGCGATGCTCTCATCACAGAGCATGTGGAGGCTTTATCA 925  
QY 611 TGATGTTGGGCTTGAGCTACTCTTTGTTGTTTCTTATCCCTTTGATGTTTCATGATATTT 670  
Db 926 CCATCCAGTGGCCAGATGGTATCGGCTTCTGTTCCCTGCTGGCCATGAGCTTCT 985  
QY 671 GTTACAGCTTCATTGTCAAAACCTTTGGTGCAGAGTTCAGAAATTTCTAAAGCGCAAGGCCA 730  
Db 986 GTTACCTTGTTCATCATCGCACCCCTGCTCCAGGCACGCAACTTTGAGCGCAACAAGGCCA 1045  
QY 731 TCCGTGTAATCATAGCTGTGTGTTGTTGTTCTGGCTTGTCTGATTCCTCTCAACA--- 787  
Db 1046 TCAAGGTGATCATCGCTGTGTGGTCTTTCATAGTCTTCCAGGCTGCCCTACAATGGGG 1105  
QY 788 TGGTCTCTGTTGTGACGGCTGCTAATTTGGGTAATAATGAACCGATCTGCCAGAGCGAAA 847  
Db 1106 TGGTCTTGCCCGCAGACGGTGGCCAACTTCAACATCACCAGTACACCTGTGAGCTCAGTA 1165  
QY 848 AGCTAATTTGGCTATACGAAACCTGTACAGAAAGTCTGCTGCTTCTGCTGCTGCTGCTGA 907  
Db 1166 AGCAACTCAACATCGCTACGACGTCACCTACAGCCTGGCCTGCTGCCGTCA 1225  
QY 908 ACCGTGCTCTAGCTTTTATTTGGCGAGAGTTTCAGAACTACTTTCTGAGAGATCTTGA 967  
Db 1226 ACCCTTCTTGTAGCGCTTTCATCGGCTCAAGTTCCGCAACGATCTCTTCAAGCTCTTCA 1285  
QY 968 AGGACCTGTGTGTGAGAGGAAGTA 995

QY	311	TCAGCAATGCCACGTCGCAAGTTGCTAAAGAGGATCTATGCCATCAACTTTAACTGCGGGA	370
DB	632	TCGGTGTCCACATTTTGCAAGCTCATCTTTGCCATCTACAAGATGAGCTTCTTCAGTGCCA	691
QY	371	TGCTGCTCTGACTTGTGCATTAGCATGACCGGTACATCGCATGTGACAGCGACTAAGT	430
DB	692	TGCTCTTACTCTTTTGATCATGATGACCGTACGTGGCATCGTCCAGGCTGTCTCAG	751
QY	431	CATTCGGGTCGGATCCAGAACACTACCGCGCAGCAAAAATCATCTGCCTTGTGTGTGGG	490
DB	752	CTCACCGGCACCGTCCCGCGCTTCTTCATCAGCAAGCTGTCTGTGTGGCATCTGGA	811
QY	491	GGCTGTGAGTCATCTCCAGCTCAACTTTTGTCTTTCACCAAAAAATACACACCCCAAG	550
DB	812	TACTAGCCACAGTGTCTCTCATCCAGAGCTCTCTGTACAGTGAACCTCCAGAGA-----	865
QY	551	GCAGCGATGCTCTGTAAACCAAGTACCAAACTCTCTCGGAGCCCATCAGTGTGAAGCTGC	610
DB	866	GCAGCAGTGAGCAAGGATGCGATGCTCTCTCATCACAGAGCATGTGGAGGCCCTTTATCA	925
QY	611	TGATGTTGGGGCTTGAGCTACTCTTTTGGTTCTTTATCCCTTTTGATCTTCATGATATTTT	670
DB	926	CCATCCAGGTGCCCCAGATGGTGATCGGGCTTTCTGTCTCCCGTCTGGCCATCAGCTTCT	985
QY	671	GTTACAGGTTCAATTGTCAAAACCTTGGTCAAGCTCAGAAATTCCTAAAGCACAAGCCA	730
DB	986	GTTACTTTGTCTATCATCGGCACCCCTGCTCCAGCGACGCAACTTTTGACGCGCAACGGCCA	1045
QY	731	TCGGTGTAAATCACTAGCTGGTGGCTTGCTTTTCTGGCTTGTCAAGATTCCTCATAACA---	787
DB	1046	TCAAGGTGATCATCGCTGTGGTGTGGTCTTCATAGCTTCACAGTGCCTCAATATGGGG	1105
QY	788	TGGTCTCTTGTGACGGCTGCTAAATTTGGGTAAAATGAACCGATCCTGCCAGAGCGAAA	847
DB	1106	TGGTCTTGCCCCAGACGGTGGCCAACTTCAACATCACCAAGTAGCACCTGTGAGCTCAGTA	1165
QY	848	AGCTAATTGGCTATACGAAAACGTGCACAGAGTCCTGCGTTTCTGTCACCTGCTGCTCGA	907
DB	1166	AGCAACTCAACATCGCCTACGACGTCAACCTACAGGCTGGCCTGCGTCCGCTGCTGGTCA	1225
QY	908	ACCCCTGTGCTCTACCGCTTTTATTTGGCGAGAAGCTTCAGAAAACACTTCTTGAAGATCTTGA	967
DB	1226	ACCCCTTCTTGTACGCCTTTATCGGGGTCAAGTTCGGCAACGATCTCTTCAAGCTCTTCA	1285
QY	968	AGGACCTGTGGTGTGTGAGAAAGGAAGTA	995
DB	1286	AGGACCTGGGCTGGCTCTCAGCCAGGAGCA	1313

```

RESULT      7
US-08-153-848-14
: Sequence 14, Application US/08153848
: Patent No. 5759804
: GENERAL INFORMATION:
: APPLICANT: Godiska, Ronald
: APPLICANT: Gray, Patrick W.
: APPLICANT: Schweikart, Vicki L.
: TITLE OF INVENTION: NO. 5759804el Seven Transmembrane Receptors
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: ADDRESSEE: Bicknell
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,848
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5759804and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2160 base pairs
TYPE: nucleic acid
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 64..1197
US-08-153-848-14

Query Match      19.8%  Score 222;  DB 1;  Length 2160;
Best Local Similarity 55.5%;  Pred. No. 7.7e-56;
Matches 515;  Conservative 0;  Mismatches 400;  Indels 13;  Gaps 4;

QY 71 GTTACTGTCCTCTTCCAGGAGTCCAGGAGTCTCTCCAGGCTATTGTACCGAATGCCT 130
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 198 GTCTTTGTCCTCAAGAGGACGTGCGAATT-TAAGCTGTTCCCTCCCTATCATGT 256
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 131 ACTCTTGATCTGTGTTTGGGCTCCTGGGGAATATCTGGTGGTGATCACTTGGCTT 190
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 257 ACTCATATTTTTCGTGGGCTACTGGGCAATGGGCTGGTGGTGGTGGTGGTGGTATCT 316
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 191 TTATAGAGGCGAGGTCTATCAGACGCTCTATCTTTGAACATGGCCATTCGACACA 250
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 317 ATTCAGAGGCTCAAGACCATGACCGATACCTTACCTGCTCAACCTGGCGGTGGCAGACA 376
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 251 TCCTCTGTTCTTACTCTCCCATCTGGSCAGTGAATCATGCACTGGTGGTGGGTTT 310
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 377 TCCTCTCTCTGACCCCTCCCTTCTGGGCTACAGCGGGCA--AGTCTGGGTCT 433
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 311 TCAGCAATGCCAGTCAAGTGTCTAAAAGGATCTATGCGCATCACTTTAACTGCGGGA 370
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 434 TCGGTGTCCACTTTTGAAGCTCATCTTTGCCATCTACAAGATGAGCTTCTTCAGTGGCA 493
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 371 TGTGTCCTGACTTCGATTAGCATGAGCGGTACATCGCATGTCATGAGCGCACTAAGT 430
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 494 TGTCTCTACTTCTTGATCAGCATTTGACGCTACGTGGCCATCGTCCAGGCTGTCTCAG 553
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 431 CATTCGGGTCGGATCCAGAACACTACCGCGACGAAATCATCTGCTGTTGTGTGGG 490
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 554 CTACCGCCACCGTGCCCGGCTCTCTCATCAGCAAGCTGTCTGTGGGGCATCTGGA 613
   || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 491 GGCTGTGATCATCATCTCCAGCTCAACTTTTGTCTTCAACCAAAAATACAACACCCAAG 550
   || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 614 TACTAGCCACAGTGTCTCCATCCCAAGAGCTCTCTGATGAGTACAGTCCAGAGGA----- 667
   || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 551 GCAGCATGTCTGTGNAACCAAGTACCAACTGTCTCGGAGGCCCATCAGTGAAGCTGC 610
   || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 668 GCAGCAGTGAAGCAGCGATGCGATGCTCTCTCATCAGACGATGTGGAGGCCCTTATCA 727
   || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 611 TGATGTTGGGCTTGAGCTACTTGTGTTTCTTTTATCTCTTGTGTTTCATCATATTTT 670
   || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 728 CCATCAGGTGCCCAAGATGGTATCGGCTTCTTGTGTCCTTGTGGCCATGAGCTTCT 787
   || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 671 GTTACAGTTTATGTCAAAACCTTGGTGCAAGCTCAGAATTTCTAAAAGGCACAAAAGCCA 730
   || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||

```

TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 64...1197  
US-09-299-843A-14

Query Match 19.8%; Score 222; DB 4; Length 2160;  
Best Local Similarity 55.5%; Pred. No. 7.7e-56;  
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

QY 71 GTTACTGTGCTCTTGGCAGGAGTCCAGGAGTCTTCCAGGCTATTGTACCGAATTGGCT 130  
Db 198 GTCTTTGTGCTCCAGAGGAGGCTGCGAACTT-TAAAGCCTGGTTCCTCCCTATATCT 256  
QY 131 ACTCCTTGATCTGTGTCTTGGGCTCTCTGGGAAATATCTGTGGTGGTATCACCTTTGCTT 190  
Db 257 ACTCCATCATTTGTTTCGTGGGCTTACTGGCAATGGCTGGTGTGACCTATATCT 316  
QY 191 TTTATAAGAGGCCAGGTCTATGACAGAGCTATCTCTTTGAACATGGCCATTGACAGACA 250  
Db 317 ATTTCAAGAGGCTCAAGACCATGACCGATACCTACCTGTCTCAACCTGGCGTGGCAGACA 376  
QY 251 TCCCTCTTTCTTCTTACTCTCCCATTTCTGGCAGTGTAGTCCACTGGTGGGTTT 310  
Db 377 TCCCTCTTCTCTGACCTTCCCTTCTGGGCTACAGCGCGGCCA---AGTCTGGGTCT 433  
QY 311 TCAGCAATGCCAGTGCAGTGTCTTAAAGGATCTATGCCATCACTTAACTCGGGA 370  
Db 434 TCGGTGCCATTTTGAAGCTCACTTTTGGCACTACAGATGAGCTTCTTCAGTGGA 493  
QY 371 TGTCTCTCTGCTTACGATTAGCATGGACCGGTATACCGCATGTGTACAGCGACTAAGT 430  
Db 494 TGTCTCTACTTCTTGTGATCAGCATGACCGGTACCGGTACCGGTCTCCAGGCTGTCTCAG 553  
QY 431 CATTCGGGCTCCGATCCAGACACTACCGCGCAGCAAAATCATCTGCTTGTGTGGG 490  
Db 554 CTCACCGGACCGGCGCGCTTCTCTCATCAAGCAAGCTGTCTGTGGGCACTCTGA 613  
QY 491 GGCTGTCAGTCAATCATCTCCAGCTCAACTTTTGTCTTCAACCAAAATACACACCCCAAG 550  
Db 614 TACTAGCCACAGTGTCTTCCATCCAGAGCTCTGTACAGTGAACCTCCAGAGA----- 567  
QY 551 GCAGCGATCTGTGTAACCCCAAGTACCAAACTGTCTCGAGGCCATCAGGTGGAAGCTGC 610  
Db 668 GCAGCAGTGACGAGCGATGCGTCTCTCATCACAGAGCATGTGGAGGCTTTATCA 727  
QY 611 TGATGTTGGGCTTGAGTACTCTTTTGTGTTCTTATCCCTTTGATGTTTCAATATTTT 670  
Db 728 CCATCCAGGTGGCCCGATGATGCTGCTTCTGGTCCCTGCTGGCCATGAGCTTCT 787  
QY 671 GTTACACGTTTCAATGTAACCACTTGGTGAAGTCAAGATTTCTAAAGGACCAAGCCA 730  
Db 788 GTTACCTGTGATCATCTCCGACCTGCTCCAGGACCGCAACTTTGAGCGCAACAGGCCA 847  
QY 731 TCCGTGTAATCATAGCTGTGTGCTTGTGTTTCTGGCTTGTGAGATTCCTCATACA--- 787  
Db 848 TCAAGGTGATCATCGCTGTGTGCTGCTTCTCATAGTCTTCCAGCTGCGCTACATGGG 907  
QY 788 TGGTCTCTGTTGAGCGCTGCTAATTTGGGTAAATGAACCGATCTCTGACAGCGGAAA 847  
Db 908 TGGTCTCTGGCCAGCGGTGGCCAACTTCAACATCACCAGTAGCACCTGTGAGCTCAGTA 967  
QY 848 AGCTAATTTGGGTATACGAAACTGTACAGAGTCTCTGGCTTCTGTGACCTGCTCCCTGA 907  
Db 968 AGCAACTCAACATCGCTTACGACGTCACTTACAGCTTGGCTGGCTGCTCGCTCGCTCA 1027  
QY 908 ACCCTGTCTCTAGCTTTTATTTGGGCAAGAGTTCAGAAACTACTTCTTGAAGATCTTGA 967  
Db 1028 ACCCTTTCTGTAGCGCTTCTATCGCGCTCAAGTTCGCGCAACGATCTCTTCAAGCTCTCA 1087  
QY 968 AGGACCTGTGTGTGAGAGGAAGTA 995

Db 1088 AGGACCTGGCTGCTCCTCAGCCAGGACA 1115

## RESULT 9

PCT-US93-11153-14  
; Sequence 14, Application PC/TUS9311153  
; GENERAL INFORMATION:  
; APPLICANT: Godiska, Ronald  
; APPLICANT: Gray, Patrick W.  
; APPLICANT: Schweikart, Vicki L.  
; TITLE OF INVENTION: Novel Seven Transmembrane Receptors  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/11153  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: US 07/977,452  
; FILING DATE: 17-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Noland, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31794  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2160 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 64...1197  
PCT-US93-11153-14

Query Match 19.8%; Score 222; DB 5; Length 2160;  
Best Local Similarity 55.5%; Pred. No. 7.7e-56;  
Matches 515; Conservative 0; Mismatches 400; Indels 13; Gaps 4;

QY 71 GTTACTGTGCTCTTCCAGGAGTCCAGGAGTCTTCCAGGCTATTGTACCGAATTGGCT 130  
Db 198 GTCTTTGTGCTCCAGAGGAGGCTGCGAACTT-TAAAGCCTGGTTCCTCCCTATATCT 256  
QY 131 ACTCCTTGATCTGTGTCTTGGGCTCTCTGGGAAATATCTGTGGTGGTATCACCTTTGCTT 190  
Db 257 ACTCCATCATTTGTTTCGTGGGCTTACTGGCAATGGCTGGTGTGACCTATATCT 316  
QY 191 TTTATAAGAGGCCAGGTCTATGACAGAGCTATCTCTTTGAACATGGCCATTGACAGACA 250  
Db 317 ATTTCAAGAGGCTCAAGACCATGACCGATACCTACCTGTCTCAACCTGGCGTGGCAGACA 376  
QY 251 TCCCTCTTTCTTCTTACTCTCCCATTTCTGGCAGTGTAGTCCACTGGTGGGTTT 310  
Db 377 TCCCTCTTCTCTGACCTTCCCTTCTGGGCTACAGCGCGGCCA---AGTCTGGGTCT 433





```
; Sequence 1, Application US/08383751A
; Patent No. 5753516
; GENERAL INFORMATION:
; APPLICANT: Heavy, Wyrta E.
; APPLICANT: Finberg, Robert W.
; TITLE OF INVENTION: Identification and Uses of Opioid
; TITLE OF INVENTION: Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: US
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,751A
; FILING DATE: 03-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: DFCI:001/WIM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2154 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64..1197
; US-08-383-751A-1

Query Match 19.3%; Score 215.6; DB 1; Length 2154;
Best Local Similarity 55.1%; Pred. No. 5.9e-54;
Matches 511; Conservative 0; Mismatches 404; Indels 13; Gaps 4;

QY 71 GTTACTGTCTCTTCCAGGAGGTGAGGAGTCTCTCCAGGCTATTGTACCGAATTGCCT 130
Db 198 GTCTTTGTCTCCAGAGGAGGTGCGGAACTT-TAAAGCCGTGGTCTCCCTATCATGT 256

QY 131 ACTCTTGATCTGTCTTGTGCTTGGCCCTCTGGGAAATATCTGTGTGTGATCACTTTGCTT 190
Db 257 ACTCATCATTTTGTGCGGCGCTTACTGGGCAATGGGCTGTGCTGTGTGACCTATATCT 316

QY 191 TTATAGAGGCCAGGCTATGACAGAGGCTCTCTTGTGACATGGCCATGGCAGACA 250
Db 317 ATTCAAGAGGCTCAAGACCATGACCGATACCTAGCTGTCTCAACCTGGCGGTGGCAGACA 376

QY 251 TCCTCTTTCTTCTTACTCTCCCATTTGCGGAGTGAGTGATGCTGCTGCTGGTGGT 310
Db 377 TCCTCTTCTCTGACCTTCTCCCTTCTGGCCCTACAGCGCGGCA---AGTCTGGGTCT 433

QY 311 TCAGCAATGCCACGTGCAAGTGTCTTAAAGGCACTATGCCATCACTTTAACTGCGGGA 370
Db 434 TCGGTGCCACTTTTGAAGCTCATCTTTGGCCATCTACAAGATGAGCTTCTCAGTGCCA 493

QY 371 TGCTGCTCTGACTTCATAGCATGAGCGGTACATGCCCATGATGACAGGCGCACTAAGT 430
Db 494 TGCTCTACTTCTTTGTCATGACCATTTGACCGCTAGTGGCCATGCTCCAGGCTGTCTCAG 553
```

```
QY 431 CATTCGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTGTTGTGTGGG 490
Db 554 CTCACCGCCACCGTGCCTGCTCTCTCATCAGCAAGCTGCTCTGTGTGGGCGAGGCCA 613

QY 491 GGCTGTGAGTCACTCTCCAGCTCAACTTTTGTCTTCAACCAAAATACACACCCCAAG 550
Db 614 TACTAGCCACAGTGTCTCTCCATCCAGAGCTCTCTGTAGAGTGAACCTCCAGAGGA----- 667

QY 551 GCAGCGATGTGTGAACCAAGTACCAAACTGTCTCGAGGCCCATCAGGTGGAAGCTGC 610
Db 668 GCAGAGTGAAGCAAGCATGCGATGCTCTCTCATCAGAGAGCATGTGGAGGCTTTATCA 727

QY 611 TGATGTTGGGCTTTGAGCTACTCTTTGTTTCTTTATCCCTTTGATGTTGATGATATTT 670
Db 728 CCATCCAGTGGCCAGATGCTGCTGCTTCTGTTCCCTGCTGGCCATGAGCTTCT 787

QY 671 GTTACAGTTCATTTGTCAGAACTTTGGTGGCAAGTTCAGAAATTTAAAAGGCAACAAGCA 730
Db 788 GTTACCTTTGTCATCATCGCACCTGCTCCAGGCAACGCACTTTGAGGCGCAACAAGGCCA 847

QY 731 TCCGTGTAATCATAGCTGTGCTGCTGTTGCTGCTTGTGCTGCTGTCAGATTCCTCATAACA--- 787
Db 848 TCAAGTGTATCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907

QY 788 TGGTCTGCTGTTGACGGCTGCTAAATTTGGGTAAATGAACGATCTCTGCCAGAGCGAAA 847
Db 908 TGGTCTGCTGCTGACGGCTGCTAAATTTGGGTAAATGAACGATCTCTGCCAGAGCGAAA 967

QY 848 AGTAAATGGCTATACGAAACTGTCACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907
Db 968 AGCAACTCAACATCGCTTACGACGCTACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027

QY 908 ACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 967
Db 1028 ACCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087

QY 968 AGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 995
Db 1088 AGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115

RESULT 13
US-08-352-678-1
; Sequence 1, Application US/08352678
; Patent No. 6043351
; GENERAL INFORMATION:
; APPLICANT: Birkenbach, Mark
; APPLICANT: Kieff, Elliott
; TITLE OF INVENTION: EPSTEIN BARR VIRUS INDUCED GENES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/352,678
; FILING DATE: 30-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/980,518
; FILING DATE: 25-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gates, Edward R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: B0801/7044
```



Db 434 TCGGTGTCACATTTTGCAGAGCTCATCTTTGCCATCTACAAGATGAGCTTCTTCAGTGGCA 493  
QY 371 TGTCTGCTCTGACTGTGCTATGATGACGACCGGTACATGCGCATGTTGACAGCGGACTAAGT 430  
Db 494 TGTCTGCTACTTCTTGTGATCAGATGACCGGTACGCTGCGCATGTTGACAGCGGCTGCTCAG 553  
QY 431 CATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCCCTTGTGTTGGG 490  
Db 554 CTCACGGCCACCGTGCCTCTCTCATCAGCAAGCTGTCTGTGTTGGGCGGCGCA 613  
QY 491 GGTGTGCTGATCATCTCTCCAGCTCAACTTTTGTCTTCAACCAAAATACACACCCCAAG 550  
Db 614 TACTAGCCACAGTGTCTCCATCCAGAGCTCTGTGACAGTACCTCCAGAGCA----- 667  
QY 551 GCAGGATGCTGTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGTGAAGCTGC 610  
Db 668 GCAGCAGTGAAGGATGCGATGCTCTCATCAGAGCATGTTGAGGCGCTTTATCA 727  
QY 611 TGATGTTGGGCTTGAGCTPACTTCTTTGGTCTTTTATCCCTTTGATGTTTCATGATATTT 670  
Db 728 CCATCCAGTGGCCAGATGGTGATCGGCTTCTGTGCTCCCTGCTGGCATGAGCTTCT 787  
QY 671 GTTACAGCTTATGTCAAAACCTGGTGCAGCTCAGATTTCTAAAGCAGCAAAAGCCA 730  
Db 788 GTTACCTTGTCTATCATCCGACCTGCTCCAGCAGCAAACTTTGAGCGCAACAAGCCA 847  
QY 731 TCCGTGTAATCATAGCTGTGTGCTGTCTTCTGCTTGTGCTGATTCCTCATAACA--- 787  
Db 848 TCAAGTGATCATGCTGTGGTGGTCTTCTCATAGCTTCTCCAGCTGCGCTTCAATGGGG 907  
QY 788 TGGTCTGCTGTGAGCGCTGTCTAATTTGGGTAAATGAACCGCATCTGCCAGAGCGAA 847  
Db 908 TGGTCTGCGCCAGCGGTGGCCAACTTCAACATCACAGTACAGCTGTGAGCTCAGTA 967  
QY 848 AGCTAATGGCTATAGAAACTGTACAGAGTCTCGCTTTCCTGCTGCTGCTGCTGA 907  
Db 968 AGCAACTCAACATCGCTACGAGCTCACCTACAGCGCTGCGCTGCTGCTGCTGCTCA 1027  
QY 908 ACCCTGTGCTCTACGCTTTTATTTGGCAGAGTTCAGAACTACTTCTCAAGATCTTGA 967  
Db 1028 ACCCTTCTGTGAGCTTCTATCGGCTCAAGTTCGCAACGATATCTTCAAGCTCTTCA 1087  
QY 968 AGGACCTGTGTGTGAGGAAGTA 995  
Db 1088 AGGACCTGGGCTGCTCAGCAGGACCA 1115

RESULT 15  
US-08-153-848-23  
Sequence 23, Application US/08153848  
Patent No. 5759804  
GENERAL INFORMATION:  
APPLICANT: Godiska, Ronald  
APPLICANT: Gray, Patrick W.  
APPLICANT: Schweikart, Vicki L.  
TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/153,848  
FILING DATE:

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,452  
FILING DATE: 17-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5759804and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31794  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2751 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1..691  
NAME/KEY: exon  
LOCATION: 692..1771  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 692..1768  
FEATURE:  
NAME/KEY: polyA\_signal  
LOCATION: 2341..2348  
US-08-153-848-23

Query Match 18.4%; Score 206.4; DB 1; Length 2751;  
Best Local Similarity 55.3%; Pred. No. 3.4e-51;  
Matches 510; Conservative 0; Mismatches 396; Indels 16; Gaps 5;

QY 71 GTTACTGTGCTTTCAGGAGGTCAGGAGTCTTCCAGGCTATTGTACCGAATTCGCT 130  
Db 766 GTCGGTGTGCTTCAAGAAGGATGTGCGGAACCTT-TAAGGCTTGGTTCCTGCTCTCATGT 824  
QY 131 ACTCCTTGTATCTGTCTTGGCCCTCCCTGGGGAATATTCTGCTGGTGATCACCTTTGCTT 190  
Db 825 ATCTGTCTATCTGTCTGCTGGCCCTGCTCGGCAAGGGCTGGTGATCTAGCTACATCT 884  
QY 191 TTTATAAGAGGCGCAGGTCTATGACAGAGCTCTATCTTTGAACATGGCCATTCGACAGA 250  
Db 885 ATTTCAAGAGGCTCAAGACCATGACGGATACCTACTGCTCAACCTGGCGTGGCAGACA 944  
QY 251 TCCCTTTGTTTCTTACTCTCCCATTTCTGGGAGTGTGATGATGCCACTGTGCTGGGTTT 310  
Db 945 TCCTTTTCTCTTAATTTCTCCCTTCTGGGCTTACAGCGAAGCCA---AGTCTGGATCT 1001  
QY 311 TCAGCATGCCACGTGCAAGTTGCTTAAAGGATCTATGCCCATCAACTTTAACTGCGGGA 370  
Db 1002 TTGGCGTCTACTGTGTAAGGATCTTTGGCATCTATTAAGTTAAGCTTCTTTCAGCGGA 1061  
QY 371 TGCTGCTCTGACTTGGATTAGCATGGACCGGTGATCATGCCATTTGTACAGGCGACTAAGT 430  
Db 1062 TGCTGCTCTCTCTATGATCATGACATTGACCGCTACGTAGCCATCGTCCAGGCGGTGCTCG 1121  
QY 431 CATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTTGTGTTGGG 490  
Db 1122 GTCATCGCCACCGCGCGCTGCTTCTCATCAGAAAGTGTCTGTGTGGCATCTGGA 1181  
QY 491 GGCTGTGATCATCTCTCCAGCTCAACTTTTGTCTTCAACCAAAATACACACCAAG 550  
Db 1182 TGCTGGGCTCTCTCTCTCTCCATCCCGGAGCTGCTCTACAGCGGCGCTCCAGAGAACAGCG 1241  
QY 551 GCAGCGATGTCTGTGAACCCAGTACCAAACTGTCTCGGAGGCCATCAGTGTGAAAGCTGC 610  
Db 1242 CCGAGGAC-----ACGCTGATGCTCTACTGCTAGTGCCCAAGTGGAGGCTTGTATCA 1295



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2001, 08:25:29 ; Search time 119.87 Seconds  
(without alignments)  
5449.664 Million cell updates/sec

Title: US-08-887-977-9

Perfect score: 1119

Sequence: 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 678276 seqs, 291890651 residues

Word size : 15

Total number of hits satisfying chosen parameters: 823

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq\_0401.\*

```

1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1068	95.4	1119	19	V15418 Human dendritic ce
2	1047	93.6	1098	21	A51971 Primate (human) ch
3	24	2.1	24	21	A47508 Primer for amplifi
4	23	2.1	34	19	V31479 PCR primer HSU4598
5	21	1.9	21	21	Z60074 Reverse PCR primer
6	21	1.9	21	21	Z60077 Probe -622/CCR6 fo
7	20	1.8	20	21	A47507 Primer for amplifi
8	19	1.7	2277	20	Z00369 Nucleotide sequenc
9	19	1.7	2277	20	Z00345 Nucleotide sequenc
10	19	1.7	3293	11	Q05753 Microspore-specifi
11	19	1.7	3293	18	T48862 Brassica napus mic

c 12	18	1.6	602	20	X20419	Human secreted pro
c 13	18	1.6	822	20	X37413	Human secreted pro
c 14	18	1.6	1092	21	C69465	Human secreted pro
c 15	18	1.6	1410	21	C98169	Human colon cancer
c 16	18	1.6	2606	20	Z08776	Hop lupulin regula
c 17	18	1.6	3439	20	Z08775	Hop lupulin regula
c 18	18	1.6	6935	21	Z93304	Sequence encoding
c 19	18	1.6	134499	21	F22286	BAC containing rep
c 20	18	1.6	1664976	19	V21209	Methanococcus jann
c 21	17	1.5	300	20	Z14869	Human gene express
c 22	17	1.5	380	21	C03809	Human secreted pro
c 23	17	1.5	478	21	A80659	Human secreted pro
c 24	17	1.5	686	19	X14062	H. pylori GHPO 454
c 25	17	1.5	756	19	X14375	H. pylori GHPO 453
c 26	17	1.5	768	20	Z15412	Human gene express
c 27	17	1.5	822	21	C36461	Arabidopsis thalia
c 28	17	1.5	861	21	C35528	Arabidopsis thalia
c 29	17	1.5	982	21	C53013	Arabidopsis thalia
c 30	17	1.5	1032	14	Q47197	DNA polymerase III
c 31	17	1.5	1032	18	T98154	E. coli DNA polyme
c 32	17	1.5	1059	18	T85163	Macaque chemokine
c 33	17	1.5	1059	19	V23992	Human CC-CKR5 codi
c 34	17	1.5	1071	20	V84123	HIV-1 co-receptor
c 35	17	1.5	1225	19	T76920	DNA encoding human
c 36	17	1.5	1255	19	T76919	DNA encoding human
c 37	17	1.5	1272	22	F32719	Human secreted pro
c 38	17	1.5	1344	20	V84159	HIV-1 co-receptor
c 39	17	1.5	1376	20	V84126	HIV-1 co-receptor
c 40	17	1.5	1414	18	T44042	Human G-protein ch
c 41	17	1.5	1414	21	Z91481	Human G-protein ch
c 42	17	1.5	1442	18	T90118	cDNA for inactive
c 43	17	1.5	1477	18	T90117	cDNA for human CCR
c 44	17	1.5	1521	21	A05519	Streptococcus pneu
c 45	17	1.5	1526	18	T80107	Human KF-1 gene cl
c 46	17	1.5	1557	18	T95542	Human chemokine re
c 47	17	1.5	1569	21	F21023	Human low adenosin
c 48	17	1.5	1569	21	A34901	Human adenosine re
c 49	17	1.5	1641	18	T80102	KF-1 gene open rea
c 50	17	1.5	1761	21	C46304	Arabidopsis thalia
c 51	17	1.5	1767	21	C42034	Arabidopsis thalia
c 52	17	1.5	2058	18	T80103	Human KF-1 gene op
c 53	17	1.5	2080	18	T80098	KF-1 gene. Homo s
c 54	17	1.5	2135	10	Q02069	Human muscarinic a
c 55	17	1.5	2135	10	N92069	Human muscarinic a
c 56	17	1.5	3383	18	T85161	Human chemokine re
c 57	17	1.5	3383	21	F21271	Human low adenosin
c 58	17	1.5	3383	21	A35149	Human adenosine re
c 59	17	1.5	3423	18	T80099	Human KF-1 gene.
c 60	17	1.5	3526	18	T80106	Human KF-1 gene cl
c 61	17	1.5	3791	18	T63575	Chicken beta-actin
c 62	17	1.5	5674	20	Z24738	Human chemokine re
c 63	17	1.5	6273	19	V52154	Streptococcus pneu
c 64	17	1.5	6707	21	F21026	Human low adenosin
c 65	17	1.5	6707	21	A34904	Human adenosine re
c 66	17	1.5	24978	20	X60209	SEQ ID 3 of US5914
c 67	17	1.5	24979	21	A52321	Genomic DNA sequen
c 68	17	1.5	44453	20	X23519	Human kidney amino
c 69	17	1.5	95223	21	F22282	BAC containing rep
c 70	17	1.5	143068	21	F21105	Human low adenosin
c 71	17	1.5	143068	21	F21272	Human low adenosin
c 72	17	1.5	143068	21	A34983	Human adenosine re
c 73	17	1.5	143068	21	A35150	Human adenosine re
c 74	17	1.5	149412	21	A35151	Human low adenosin
c 75	17	1.5	152740	21	F21273	Human low adenosin
c 76	16	1.4	60	13	Q23553	Human CF intron 1
c 77	16	1.4	180	21	C08708	Human secreted pro
c 78	16	1.4	213	20	X88587	Human chromosome 1
c 79	16	1.4	276	20	V87278	EST clone B0538.
c 80	16	1.4	278	21	C32058	Human secreted pro
c 81	16	1.4	292	20	X41523	Human secreted pro
c 82	16	1.4	300	20	Z14991	Human gene express
c 83	16	1.4	300	21	A00412	Human colon cancer
c 84	16	1.4	308	20	V90618	Nucleotide sequenc

c 85	16	1.4	310	21	A78117	CDNA encoding huma	c 158	16	1.4	1489	21	C42025	Arabidopsis thalia
c 86	16	1.4	340	20	T89063	EST clone CB98. H	c 159	16	1.4	1507	20	X84446	Mouse brain CNG-3
c 87	16	1.4	367	18	T86286	Coding sequence fo	c 160	16	1.4	1508	20	X35920	Streptococcus pneu
c 88	16	1.4	410	21	C03048	Human secreted pro	c 161	16	1.4	1545	21	F12974	Aspergillus oryzae
c 89	16	1.4	413	21	A31531	Plant microsatelli	c 162	16	1.4	1551	21	C36754	Arabidopsis thalia
c 90	16	1.4	428	19	V30924	Human secreted pro	c 163	16	1.4	1654	18	T93210	Human neurotactin
c 91	16	1.4	434	21	C30788	Human secreted pro	c 164	16	1.4	1654	18	T64943	CDNA encoding a no
c 92	16	1.4	437	20	V64841	Mouse developing l	c 165	16	1.4	1654	21	A08500	Human neurotactin
c 93	16	1.4	439	21	A31529	Plant microsatelli	c 166	16	1.4	1654	21	Z58929	Human neurotactin
c 94	16	1.4	445	21	A31701	Plant microsatelli	c 167	16	1.4	1712	21	C43009	Arabidopsis thalia
c 95	16	1.4	477	21	C54895	Arabidopsis thalia	c 168	16	1.4	1713	21	A75152	CDNA encoding a mu
c 96	16	1.4	526	21	C45585	Arabidopsis thalia	c 169	16	1.4	1713	21	A75166	CDNA clone encodin
c 97	16	1.4	541	21	C41398	Arabidopsis thalia	c 170	16	1.4	1713	21	A75167	CDNA clone encodin
c 98	16	1.4	549	21	C93648	Cat flea head and	c 171	16	1.4	1713	21	A75168	CDNA clone encodin
c 99	16	1.4	585	21	C45921	Arabidopsis thalia	c 172	16	1.4	1809	18	T77791	DNA encoding 6-hyd
c 100	16	1.4	601	21	C40754	Arabidopsis thalia	c 173	16	1.4	1836	13	Q26053	Glucosylase cDNA
c 101	16	1.4	606	20	Z21473	Triticum aestivum	c 174	16	1.4	1839	21	C46455	Arabidopsis thalia
c 102	16	1.4	624	21	A16504	Human colon cancer	c 175	16	1.4	2008	21	C48800	Arabidopsis thalia
c 103	16	1.4	651	21	C44918	Arabidopsis thalia	c 176	16	1.4	2033	21	C36161	Arabidopsis thalia
c 104	16	1.4	659	21	C46017	Arabidopsis thalia	c 177	16	1.4	2032	21	F21251	Human low adenosin
c 105	16	1.4	662	21	C32723	Arabidopsis thalia	c 178	16	1.4	2022	21	A35129	Human adenosine re
c 106	16	1.4	664	21	A52634	Eosinophil activat	c 179	16	1.4	2097	21	Z39106	Human testin (HTE
c 107	16	1.4	670	21	A02046	Human colon cancer	c 180	16	1.4	2132	21	C36443	Arabidopsis thalia
c 108	16	1.4	676	21	C47722	Arabidopsis thalia	c 181	16	1.4	2134	21	C48799	Arabidopsis thalia
c 109	16	1.4	714	21	F14116	Aspergillus oryzae	c 182	16	1.4	2166	19	Z96324	S. pneumoniae deri
c 110	16	1.4	734	20	Z16445	Human gene express	c 183	16	1.4	2226	21	C43120	Arabidopsis thalia
c 111	16	1.4	738	21	A02249	Human colon cancer	c 184	16	1.4	2247	21	A10406	Murine cardiac ank
c 112	16	1.4	767	20	V90652	Nucleotide sequenc	c 185	16	1.4	2393	20	F21253	Human secreted pro
c 113	16	1.4	773	21	A16528	Human colon cancer	c 186	16	1.4	2484	21	F21253	Human low adenosin
c 114	16	1.4	777	17	T34690	Penicillin-resista	c 187	16	1.4	2484	21	A35131	Human adenosine re
c 115	16	1.4	874	20	X40050	Prostate cancer as	c 188	16	1.4	2497	21	F21254	Human low adenosin
c 116	16	1.4	879	21	C45814	Arabidopsis thalia	c 189	16	1.4	2499	10	N90366	Human adenosine re
c 117	16	1.4	892	18	T61423	Human cytidine dea	c 190	16	1.4	2500	11	Q05784	Synthetic human 5-
c 118	16	1.4	930	20	V90852	Nucleotide sequenc	c 191	16	1.4	2500	11	Q05784	Sequence encoding
c 119	16	1.4	960	15	Q55090	fbp2 cDNA for reve	c 192	16	1.4	2500	21	A35130	Human low adenosin
c 120	16	1.4	974	21	F21745	Human breast and o	c 193	16	1.4	2500	21	A35130	Human low adenosin
c 121	16	1.4	989	21	C69539	Human secreted pro	c 194	16	1.4	2520	21	A78420	Human secreted pro
c 122	16	1.4	1026	19	V69612	Human secreted pro	c 195	16	1.4	2574	8	N70805	Sequence of the Bg
c 123	16	1.4	1035	21	C41930	Arabidopsis thalia	c 196	16	1.4	2575	8	N70804	Sequence of the Bg
c 124	16	1.4	1037	21	C42618	Arabidopsis thalia	c 197	16	1.4	2577	6	N50014	Sequence encoding
c 125	16	1.4	1053	20	X86032	Nucleic acid encod	c 198	16	1.4	2577	14	Q37481	Elmeria tenella TA
c 126	16	1.4	1053	20	X86033	Membrane penetrati	c 199	16	1.4	2599	21	C64367	Human vesicle asso
c 127	16	1.4	1055	21	C44244	Arabidopsis thalia	c 200	16	1.4	2638	11	Q03741	Human myb related
c 128	16	1.4	1128	21	C51136	Arabidopsis thalia	c 201	16	1.4	2638	19	V20468	Human B-myb oncoge
c 129	16	1.4	1130	21	C36918	Arabidopsis thalia	c 202	16	1.4	2644	21	Z33326	Human secreted pro
c 130	16	1.4	1132	21	F79214	Human lung tumour-	c 203	16	1.4	2668	20	Z07196	Human lung tumour
c 131	16	1.4	1160	19	T98743	DNA encoding a S.	c 204	16	1.4	2668	21	C79150	Human lung tumour-
c 132	16	1.4	1160	21	F13574	Aspergillus oryzae	c 205	16	1.4	2686	13	Q26054	Encodes glucoamyla
c 133	16	1.4	1192	21	F21893	Human breast and o	c 206	16	1.4	2720	9	N81657	HVT TK gene. Herp
c 134	16	1.4	1248	21	F13127	Aspergillus oryzae	c 207	16	1.4	2738	16	Q89342	Human mgluR4 cDNA.
c 135	16	1.4	1285	21	Z32663	Murine cytotoxic T	c 208	16	1.4	2764	21	C76328	Human ORFX ORF1883
c 136	16	1.4	1286	18	T46177	Cytotoxic T lympho	c 209	16	1.4	3505	11	Q03818	Turkey herpes viru
c 137	16	1.4	1295	10	N92737	Insert of clone pH	c 210	16	1.4	3664	21	A96689	CDNA encoding a mu
c 138	16	1.4	1298	21	Z56701	Human transmembran	c 211	16	1.4	3690	21	F15891	Human prostate can
c 139	16	1.4	1320	21	C33406	Arabidopsis thalia	c 212	16	1.4	3803	20	X76396	Human secreted pro
c 140	16	1.4	1383	11	Q03809	Thymidine kinase (	c 213	16	1.4	3843	21	Z93302	Sequence encoding
c 141	16	1.4	1397	21	Z29145	Soybean phytoene s	c 214	16	1.4	3880	16	T03888	Human ORFX ORF1493
c 142	16	1.4	1455	20	Z35921	Streptococcus pneu	c 215	16	1.4	4547	21	C75938	Human mgluR4 gene.
c 143	16	1.4	1455	20	Z35922	Streptococcus pneu	c 216	16	1.4	5077	22	F23926	Human secreted pro
c 144	16	1.4	1455	20	Z35923	Streptococcus pneu	c 217	16	1.4	5360	20	X13155	Enterococcus faeca
c 145	16	1.4	1455	20	Z35924	Streptococcus pneu	c 218	16	1.4	5921	20	Z32700	Human glycophospha
c 146	16	1.4	1455	20	Z35926	Streptococcus pneu	c 219	16	1.4	6132	21	C79584	Virulence gene #4.
c 147	16	1.4	1455	20	Z35928	Streptococcus pneu	c 220	16	1.4	6405	19	V52240	Streptococcus pneu
c 148	16	1.4	1455	20	Z35929	Streptococcus pneu	c 221	16	1.4	7400	21	A29863	Swine vesicular di
c 149	16	1.4	1455	20	Z35930	Streptococcus pneu	c 222	16	1.4	7400	21	Z98717	Swine vesicular di
c 150	16	1.4	1455	20	Z35932	Streptococcus pneu	c 223	16	1.4	7400	21	Z98718	Swine vesicular di
c 151	16	1.4	1455	20	Z35933	Streptococcus pneu	c 224	16	1.4	7400	21	Z98719	Swine vesicular di
c 152	16	1.4	1455	20	Z35934	Streptococcus pneu	c 225	16	1.4	7421	21	Z98720	Swine vesicular di
c 153	16	1.4	1455	20	Z35935	Streptococcus pneu	c 226	16	1.4	7585	19	V68403	Human BAZ2-beta cD
c 154	16	1.4	1455	20	Z35937	Streptococcus pneu	c 227	16	1.4	8561	20	X06982	Candida albicans h
c 155	16	1.4	1455	20	Z35938	Streptococcus pneu	c 228	16	1.4	8625	21	C75834	Human ORFX ORF1389
c 156	16	1.4	1475	19	V43795	Rodent chemokine r	c 229	16	1.4	8937	13	Q20602	Nf1 gene. Homo sa
c 157	16	1.4	1485	21	C51141	Arabidopsis thalia	c 230	16	1.4	10706	18	T46941	Human neurofibroma

231	1.4	11151	21	A35133	Human adenosine re	c 304	15	1.3	404	12	Q12906
232	1.4	11181	21	F21255	Human low adenosin	c 305	15	1.3	406	21	C28917
233	1.4	11220	21	C48974	Arabidopsis thalia	c 306	15	1.3	413	21	C19654
234	1.4	13808	20	X13208	Enterococcus faeca	c 307	15	1.3	418	20	V90289
235	1.4	29544	17	T12460	Ovine adenovirus 9	c 308	15	1.3	419	21	C31630
236	1.4	35100	19	V20441	Human c-fms oncoge	c 309	15	1.3	426	20	X08739
237	1.4	35414	21	D00147	TR12 related DNA-1	c 310	15	1.3	427	20	X41214
238	1.4	90050	21	T91925	Wild type (C57BL/6	c 311	15	1.3	433	20	X40584
239	1.4	122186	22	C89560	Human histone deac	c 312	15	1.3	436	14	Q39702
240	1.4	125910	21	C64370	Human KCNQ5 (KCN6q	c 313	15	1.3	446	14	Q59114
241	1.4	162450	21	Z86967	Retinoblastoma bin	c 314	15	1.3	447	20	V90439
242	1.4	235033	19	V57926	Hereditary haemoch	c 315	15	1.3	453	21	X41002
243	1.4	237326	19	V57903	Hereditary haemoch	c 316	15	1.3	455	20	C13913
244	1.3	20	20	Z25788	Human p51 PCR prim	c 317	15	1.3	456	21	C45718
245	1.3	28	20	X90631	Forward primer 16K	c 318	15	1.3	456	21	C03986
246	1.3	32	22	C68040	Adaptor #7 used in	c 319	15	1.3	457	21	C54820
247	1.3	38	18	V05323	Human Col III gene	c 320	15	1.3	463	21	A61339
248	1.3	48	17	T07129	Human LAP 48-mer p	c 321	15	1.3	468	21	C09121
249	1.3	54	22	C68041	Oligonucleotide #1	c 322	15	1.3	469	21	C79476
250	1.3	80	16	T24568	Human gene signatu	c 323	15	1.3	471	20	X97675
251	1.3	96	21	C14893	Human secreted pro	c 324	15	1.3	479	21	A78185
252	1.3	115	16	T24585	Human gene signatu	c 325	15	1.3	488	18	X30883
253	1.3	118	14	Q39822	Expressed Sequence	c 326	15	1.3	489	20	X13914
254	1.3	123	18	Q59234	Human brain.Expres	c 327	15	1.3	494	17	T59835
255	1.3	128	21	C15458	Human secreted pro	c 328	15	1.3	523	20	X20869
256	1.3	127	17	T07133	Lingual antimicrob	c 329	15	1.3	531	21	C03872
257	1.3	128	21	C72592	Single nucleotide	c 330	15	1.3	534	21	A02622
258	1.3	133	17	T07134	Pro-LAP coding seq	c 331	15	1.3	535	19	V20086
259	1.3	146	21	C21826	Human secreted pro	c 332	15	1.3	536	21	A44980
260	1.3	149	20	X34711	PCR amplified frag	c 333	15	1.3	547	21	A59493
261	1.3	150	21	C13406	Human secreted pro	c 334	15	1.3	549	21	C93809
262	1.3	161	20	X34708	Human secreted pro	c 335	15	1.3	549	21	C94873
263	1.3	163	21	C22189	DNA sequence of mo	c 336	15	1.3	556	20	X85719
264	1.3	166	21	C30823	Human secreted pro	c 337	15	1.3	556	20	X20451
265	1.3	187	21	C10381	Human secreted pro	c 338	15	1.3	564	19	X14346
266	1.3	221	21	Z95010	CSG Prol14 EST use	c 339	15	1.3	566	21	C52240
267	1.3	227	21	C17823	Human secreted pro	c 340	15	1.3	566	21	C54690
268	1.3	241	22	F17895	Human breast cance	c 341	15	1.3	568	20	V88522
269	1.3	258	20	V83961	Bacterial artifici	c 342	15	1.3	568	21	C79779
270	1.3	262	21	A00795	Human colon cancer	c 343	15	1.3	581	20	X00674
271	1.3	272	21	C02097	Human secreted pro	c 344	15	1.3	581	21	F07668
272	1.3	281	21	A87555	Rat hepatocyte car	c 345	15	1.3	591	19	V07189
273	1.3	286	21	C00387	Human secreted pro	c 346	15	1.3	591	20	V90286
274	1.3	289	21	C16157	Human secreted pro	c 347	15	1.3	603	21	Z80740
275	1.3	294	20	X81245	Human oligophrenin	c 348	15	1.3	606	21	C55023
276	1.3	297	19	V15588	Human HPK-1A C4.8	c 349	15	1.3	609	21	A16223
277	1.3	298	21	C21796	Human secreted pro	c 350	15	1.3	615	21	F08958
278	1.3	299	21	C21034	Human secreted pro	c 351	15	1.3	626	21	C52472
279	1.3	300	20	Z14165	Human gene express	c 352	15	1.3	632	21	F11397
280	1.3	309	21	A42072	Human secreted exp	c 353	15	1.3	644	21	A02476
281	1.3	316	21	C24174	Human secreted pro	c 354	15	1.3	650	21	C34114
282	1.3	319	21	C02957	Human secreted pro	c 355	15	1.3	653	16	Q99783
283	1.3	322	21	C14767	Human secreted pro	c 356	15	1.3	653	20	V62799
284	1.3	326	21	C27619	Human secreted pro	c 357	15	1.3	653	20	V81683
285	1.3	332	21	C03540	Human secreted pro	c 358	15	1.3	666	17	T36496
286	1.3	334	21	C29403	Human secreted pro	c 359	15	1.3	666	21	F12790
287	1.3	340	13	Q31044	HCV-1 genotype GII	c 360	15	1.3	668	22	F22442
288	1.3	340	17	T27968	Hepatitis C virus	c 361	15	1.3	674	21	F13506
289	1.3	340	21	C27521	Human secreted pro	c 362	15	1.3	676	21	F12612
290	1.3	350	17	T07135	Prepro-LAP coding	c 363	15	1.3	678	21	F18278
291	1.3	350	18	T69400	Murine metastatic	c 364	15	1.3	685	21	F12922
292	1.3	352	21	C99970	Human secreted pro	c 365	15	1.3	687	21	A74257
293	1.3	354	15	Q55233	Grapevine ribosoma	c 366	15	1.3	689	21	F13862
294	1.3	359	21	C30009	Human secreted pro	c 367	15	1.3	691	21	F14724
295	1.3	366	20	X27252	Mouse cClqR bindin	c 368	15	1.3	704	20	X03005
296	1.3	366	20	X27253	Rat cClqR binding	c 369	15	1.3	714	21	C71674
297	1.3	382	20	X51959	Human secreted pro	c 370	15	1.3	714	21	C71683
298	1.3	382	21	C03983	Human secreted pro	c 371	15	1.3	714	21	C71698
299	1.3	390	21	A57016	Human colon cancer	c 372	15	1.3	715	21	F12819
300	1.3	394	21	C19850	Human secreted pro	c 373	15	1.3	723	21	F09892
301	1.3	396	19	V20077	DNA encoding murin	c 374	15	1.3	726	20	V88865
302	1.3	401	20	X84200	DNA encoding human	c 375	15	1.3	729	19	V61147
303	1.3	401	21	C79429	CDNA sequence of h	c 376	15	1.3	729	19	V58490

VNTR locus D18S17  
Human secreted pro  
Human secreted pro  
EST clone DK70. H  
Human secreted pro  
Nucleotide encodin  
Human secreted pro  
Human secreted pro  
Expressed Sequence  
Human brain Expres  
EST clone D0440.  
Zea mays DNA fragm  
Enterococcus faeca  
Arabidopsis thalia  
Human secreted pro  
Arabidopsis thalia  
Part of the human  
Human secreted pro  
5' end of cDNA seq  
Extended human sec  
cDNA encoding huma  
Streptococcus pneu  
Enterococcus faeca  
Beta-casein gene f  
Polynucleotide seq  
Human secreted pro  
Human colon cancer  
Consensus DNA sequ  
Human secreted exp  
3' sequence of DNA  
Cat flea hindgut a  
Novel cDNA sequenc  
Human secreted pro  
H. pylori GHPO 271  
Arabidopsis thalia  
Arabidopsis thalia  
EST clone GM196.  
Human secreted pro  
Fusarium venenatum  
cDNA encoding a hu  
EST clone DJ238.  
Human colon cancer  
Arabidopsis thalia  
Human colon cancer  
Fusarium venenatum  
Arabidopsis thalia  
Aspergillus niger  
Aspergillus niger  
Human colon cancer  
Arabidopsis thalia  
Plant SAR gene pl.  
Tobacco SAR CHX in  
Tobacco protein-sy  
Maize glutathione-  
Aspergillus oryzae  
Human breast cance  
Aspergillus oryzae  
Aspergillus oryzae  
Lung cancer associ  
Aspergillus oryzae  
Loblolly pine SSR  
Aspergillus oryzae  
Aspergillus oryzae  
Human IL-1ra BAC c  
Single nucleotide  
Single nucleotide  
Single nucleotide  
Aspergillus oryzae  
Fusarium venenatum  
EST clone IA21. H  
5' cDNA sequence o  
5' fragment of pro

377	15	1.3	729	21	A06253	Human immunogenic	450	15	1.3	1188	21	C36150	Arabidopsis thalia
378	15	1.3	740	19	V61151	5' CDNA sequence o	451	15	1.3	1191	21	C35349	Arabidopsis thalia
379	15	1.3	740	19	V58494	5' fragment of pro	452	15	1.3	1201	21	C44975	Arabidopsis thalia
380	15	1.3	740	20	A06257	Human immunogenic	453	15	1.3	1205	21	C33121	Arabidopsis thalia
381	15	1.3	745	20	215529	Human gene express	454	15	1.3	1213	20	X97657	Extended human sec
382	15	1.3	746	21	C52799	Arabidopsis thalia	c 455	15	1.3	1217	20	242220	Human normal bladd
383	15	1.3	757	21	F14509	Aspergillus oryzae	c 456	15	1.3	1221	20	A52637	Eosinophil activat
384	15	1.3	761	20	217403	Human gene express	c 457	15	1.3	1231	20	X22754	Human ESRL1 DNA..
385	15	1.3	785	20	C76954	Human ORFX ORF2509	458	15	1.3	1233	22	F23929	Human secreted pro
386	15	1.3	807	21	C75888	Human ORFX ORF1443	459	15	1.3	1237	21	C44775	Zea mays DNA fragm
387	15	1.3	824	21	C42473	Arabidopsis thalia	c 460	15	1.3	1242	21	C33743	Arabidopsis thalia
388	15	1.3	832	9	N80115	Human prolactin	c 461	15	1.3	1243	21	C51580	Arabidopsis thalia
389	15	1.3	834	18	T74333	CDNA encoding huma	462	15	1.3	1244	20	Z20505	DNAX interferon-li
390	15	1.3	834	21	A05554	Streptococcus pneu	c 463	15	1.3	1244	21	C38049	Arabidopsis thalia
391	15	1.3	871	19	V12741	RBE1 transcription	c 464	15	1.3	1244	21	C50859	Arabidopsis thalia
392	15	1.3	887	21	C33277	Arabidopsis thalia	c 465	15	1.3	1246	21	C51062	Arabidopsis thalia
393	15	1.3	890	19	V34291	Human secreted pro	c 466	15	1.3	1249	21	C33981	Arabidopsis thalia
394	15	1.3	898	22	F22711	Human gastric canc	c 467	15	1.3	1249	21	C51061	Arabidopsis thalia
395	15	1.3	900	22	F22713	Human gastric canc	468	15	1.3	1288	19	V48113	Arabidopsis thalia
396	15	1.3	903	11	Q03320	CDNA of Elmeria te	469	15	1.3	1289	19	V61202	Nucleotide sequenc
397	15	1.3	904	19	296470	S. pneumoniae deri	470	15	1.3	1289	19	V58587	Full length CDNA s
398	15	1.3	906	21	C76864	Human ORFX ORF2419	471	15	1.3	1290	21	A06350	Prostate tumour sp
399	15	1.3	917	21	C95082	Cat flea hindgut a	472	15	1.3	1291	21	C40824	Human immunogenic
400	15	1.3	921	21	C45812	Arabidopsis thalia	473	15	1.3	1307	20	X25488	Arabidopsis thalia
401	15	1.3	933	19	V54014	Nucleotide sequenc	474	15	1.3	1317	21	C9457	Human cyclic nucle
402	15	1.3	936	17	T13396	Hepatocyte prolife	475	15	1.3	1320	22	C88168	Arabidopsis thalia
403	15	1.3	936	17	T13397	Hepatocyte prolife	c 476	15	1.3	1320	22	C88169	Infectious salmon
404	15	1.3	960	11	Q03293	Recombinant DNA en	477	15	1.3	1320	22	F33895	Infectious salmon
405	15	1.3	964	21	C35027	Arabidopsis thalia	478	15	1.3	1324	19	V34157	Human secreted pro
406	15	1.3	965	21	A69660	Pinus radiata NADP	479	15	1.3	1327	21	F21642	Human secreted pro
407	15	1.3	987	21	A95473	E. coli essential	c 480	15	1.3	1335	21	C42868	Human breast and o
408	15	1.3	987	21	A88700	E. coli FUN essent	481	15	1.3	1342	21	C38803	Arabidopsis thalia
409	15	1.3	1001	21	C57686	Arachidonic acid m	482	15	1.3	1343	21	C49751	Arabidopsis thalia
410	15	1.3	1029	15	O65614	Bovine zona pelluc	c 483	15	1.3	1374	21	A39407	Corn SYR2 homologu
411	15	1.3	1029	20	Z22719	Bovine zona pelluc	484	15	1.3	1376	19	V34157	Human secreted pro
412	15	1.3	1029	20	V64796	Bovine ZPB CDNA.	c 485	15	1.3	1379	21	C38566	Human secreted pro
413	15	1.3	1029	21	295656	Bovine zona pelluc	c 486	15	1.3	1380	21	A26697	Arabidopsis thalia
414	15	1.3	1029	21	246264	Bovine oocyte zona	c 487	15	1.3	1383	20	X56833	Candida albicans p
415	15	1.3	1029	21	233253	Bovine zona pelluc	c 488	15	1.3	1388	21	C44356	Chicken phdtkk-3 cd
416	15	1.3	1029	21	237808	Bovine zona pelluc	c 489	15	1.3	1390	21	C79804	Arabidopsis thalia
417	15	1.3	1031	21	C43756	Zea mays DNA fragm	490	15	1.3	1392	21	C48741	Human secreted pro
418	15	1.3	1040	10	N90048	Neisseria gonorrho	491	15	1.3	1394	21	C51240	Arabidopsis thalia
419	15	1.3	1040	13	Q29138	PIA gene. Synthet	492	15	1.3	1396	21	C39471	Arabidopsis thalia
420	15	1.3	1053	20	290529	Mouse BGCRr protei	493	15	1.3	1402	20	Z31790	Human G-alpha-13 c
421	15	1.3	1069	21	253163	Neisseria gonorrhe	494	15	1.3	1411	21	A26308	Human secreted pro
422	15	1.3	1071	20	X41975	Human myometrium t	c 495	15	1.3	1417	21	C77807	Human cancer assoc
423	15	1.3	1080	20	X08745	Nucleotide encodin	496	15	1.3	1418	21	F16042	Human prostate can
424	15	1.3	1080	10	N91259	Nucleotide sequenc	c 497	15	1.3	1419	21	F15876	Human prostate can
425	15	1.3	1091	21	C76819	Human ORFX ORF2374	c 498	15	1.3	1422	20	X35720	CDNA encoding a pr
426	15	1.3	1093	15	O77819	Human hepatic pare	c 499	15	1.3	1428	21	C33940	Arabidopsis thalia
427	15	1.3	1100	16	Q96139	Prolactin CDNA. H	500	15	1.3	1429	18	T74321	Human P2Y4 recepto
428	15	1.3	1101	19	V37229	Mouse CD14 coding	501	15	1.3	1461	21	Z65318	Human secreted pro
429	15	1.3	1101	21	Z36820	DNA encoding murin	502	15	1.3	1473	16	O87724	Human secreted cy
430	15	1.3	1119	16	O83773	Porin-5(A/B) of pu	503	15	1.3	1473	16	Q87725	Human auxillary cy
431	15	1.3	1123	21	C39641	Arabidopsis thalia	504	15	1.3	1473	16	O87726	Human auxillary cy
432	15	1.3	1125	21	A97356	Human colorectal c	505	15	1.3	1473	17	T28390	Human cytochrome p
433	15	1.3	1128	21	A79542	Pinus radiata cell	506	15	1.3	1473	17	T28391	Human cytochrome p
434	15	1.3	1138	21	C37772	Arabidopsis thalia	507	15	1.3	1473	17	T28392	Human cytochrome p
435	15	1.3	1145	21	C45592	Arabidopsis thalia	508	15	1.3	1473	17	T17394	Human derived cyto
436	15	1.3	1146	21	C59986	Human secreted pro	509	15	1.3	1481	17	T33247	RAR-alpha RNA sequ
437	15	1.3	1148	21	C32774	Arabidopsis thalia	510	15	1.3	1481	17	T33260	RAR-alpha DNA sequ
438	15	1.3	1149	21	C62312	Contig HEPATOX10.	511	15	1.3	1484	21	C77846	Human cancer assoc
439	15	1.3	1149	21	C35331	Arabidopsis thalia	512	15	1.3	1485	20	X15123	CDNA encoding huma
440	15	1.3	1150	21	C32807	Arabidopsis thalia	513	15	1.3	1486	21	C38414	Arabidopsis thalia
441	15	1.3	1150	21	C34731	Arabidopsis thalia	514	15	1.3	1498	21	C45265	Arabidopsis thalia
442	15	1.3	1151	21	C40645	Arabidopsis thalia	515	15	1.3	1507	18	T44067	Human Rasp-1 gene
443	15	1.3	1153	21	C40646	Arabidopsis thalia	c 516	15	1.3	1512	18	T59845	TSEP-1 coding sequ
444	15	1.3	1155	21	C40647	Arabidopsis thalia	517	15	1.3	1513	21	C45527	Arabidopsis thalia
445	15	1.3	1156	21	C33495	Arabidopsis thalia	518	15	1.3	1514	21	C38996	Arabidopsis thalia
446	15	1.3	1158	21	C41096	Arabidopsis thalia	519	15	1.3	1520	20	Z24901	Human secreted pro
447	15	1.3	1178	21	C47568	Arabidopsis thalia	c 520	15	1.3	1530	21	C49247	Arabidopsis thalia
448	15	1.3	1179	21	C34664	Arabidopsis thalia	c 521	15	1.3	1532	21	C40452	Arabidopsis thalia
449	15	1.3	1182	21	A75908	CDNA encoding a ch	522	15	1.3	1532	21	C49529	Arabidopsis thalia

523	1.3	1533	21	Z95011	Cancer specific ge	c 596	15	1.3	2052	21	A37961	Carrot promoter se
524	1.3	1536	21	C42454	Arabidopsis thalia	c 597	15	1.3	2056	21	A37964	Plasmid #3 DNA seq
525	1.3	1540	9	N80917	Sequence of the 3'	c 598	15	1.3	2071	20	X97960	Human #3 DNA seq
526	1.3	1542	20	V72588	Streptococcus pneu	599	15	1.3	2080	20	V74264	Human heart muscle
527	1.3	1554	21	C46367	Arabidopsis thalia	600	15	1.3	2085	19	V04262	Restin truncated c
528	1.3	1558	21	C47873	Arabidopsis thalia	c 601	15	1.3	2090	19	V68588	Nucleotide sequenc
529	1.3	1560	20	X00225	Arabidopsis thalia	c 602	15	1.3	2090	19	V38564	Human nuclear prot
530	1.3	1581	20	X55495	Clostridium chauvo	c 603	15	1.3	2090	21	Z97041	Human secreted pro
531	1.3	1593	17	T29199	Soybean Crml homol	604	15	1.3	2100	11	Q05676	Human secreted pro
532	1.3	1593	17	T77094	Rat acetylglucosam	c 605	15	1.3	2102	21	C76964	v-PA gamma. Desmo
533	1.3	1603	21	C42470	Human N-acetylgluc	c 606	15	1.3	2102	21	C69096	Human ORFX ORF2519
534	1.3	1604	19	V04252	Arabidopsis thalia	c 607	15	1.3	2109	21	Z98125	Human secreted pro
535	1.3	1605	19	V30916	Wild tomato acyltr	c 608	15	1.3	2118	22	C89166	Human signal pepti
536	1.3	1613	20	X35721	Human secreted pro	c 609	15	1.3	2132	20	V82077	Human p-HYDE 40 co
537	1.3	1620	11	Q00543	cDNA encoding a pr	610	15	1.3	2134	21	Z88156	DNA encoding a gid
538	1.3	1643	21	C39906	Vampire bat plasm	611	15	1.3	2137	11	Q05675	Human testis speci
539	1.3	1650	18	T91738	Arabidopsis thalia	612	15	1.3	2152	15	T40732	v-PA_beta. Desmod
540	1.3	1657	21	Z65017	JC virus DNA fragm	613	15	1.3	2152	15	T40732	TATA-binding prote
541	1.3	1668	20	V72116	Membrane-bound pro	614	15	1.3	2152	17	T42218	Human TATA-binding
542	1.3	1670	18	T43834	Mouse FAST-1 codin	c 615	15	1.3	2164	21	T79594	TATA-binding prote
543	1.3	1670	18	T43835	Helicobacter pylor	c 616	15	1.3	2174	18	V74513	Human ORFX ORF1208
544	1.3	1670	19	V34642	Helicobacter pylor	c 617	15	1.3	2185	21	A28932	Staphylococcus aur
545	1.3	1670	19	V34642	Helicobacter pylor	c 618	15	1.3	2218	18	T62417	Essential staphylo
546	1.3	1670	20	Z32136	Helicobacter pylor	c 619	15	1.3	2232	16	Q08835	S. anglica phospho
547	1.3	1670	20	Z32136	Helicobacter pylor	c 620	15	1.3	2232	16	Q08835	Human death associ
548	1.3	1670	21	D02019	Helicobacter pylor	c 621	15	1.3	2247	15	Q57973	DNA sequence encod
549	1.3	1670	21	D02020	Helicobacter pylor	c 622	15	1.3	2259	17	T30686	Human glycosyltran
550	1.3	1671	21	C40477	Helicobacter pylor	c 623	15	1.3	2259	17	T30686	Kaposi's sarcoma a
551	1.3	1674	21	C33164	Arabidopsis thalia	c 624	15	1.3	2268	20	V74265	Kaposi's sarcoma a
552	1.3	1679	21	F16104	Arabidopsis thalia	c 625	15	1.3	2268	20	V74265	Human heart muscle
553	1.3	1702	21	A37472	Human prostate can	c 626	15	1.3	2328	21	C77759	Human heart muscle
554	1.3	1704	10	N90725	DNA encoding jojob	c 627	15	1.3	2350	21	Z33368	Human cancer assoc
555	1.3	1713	17	T33900	Groundnut stilbene	c 628	15	1.3	2359	19	V52376	Streptococcus pneu
556	1.3	1713	17	T33900	Human G-protein co	c 629	15	1.3	2389	21	A79263	Pinus radiata cell
557	1.3	1713	21	A70768	Human G-protein co	c 630	15	1.3	2408	17	T30737	Rat cryptidn 2 gen
558	1.3	1713	21	Z60116	Human G-protein co	c 631	15	1.3	2418	21	F16284	Human prostate can
559	1.3	1714	20	Z33963	Human PRO617 nucle	c 632	15	1.3	2457	17	T30736	Rat cryptidn 1 gen
560	1.3	1714	21	C78478	Human PRO617 (UNQ3	c 633	15	1.3	2469	20	X13540	Enterococcus faeca
561	1.3	1714	21	A88518	Human PRO617 cDNA	c 634	15	1.3	2507	22	C86398	SNEXN DNA #2. Hom
562	1.3	1725	20	X58270	Zea mays SCLBr pro	c 635	15	1.3	2546	15	Q64889	CDNA encoding huma
563	1.3	1726	21	C48725	Arabidopsis thalia	c 636	15	1.3	2551	17	T30738	Rat cryptidn 3 gen
564	1.3	1727	21	C66225	Human cDNA encodin	c 637	15	1.3	2568	18	T62842	Rat oestrogen rece
565	1.3	1743	21	C95522	Human secreted pro	c 638	15	1.3	2568	20	X26685	DNA encoding an es
566	1.3	1763	21	C41999	Arabidopsis thalia	c 639	15	1.3	2607	13	Q24064	Encodes T3 RNA pol
567	1.3	1769	20	Z11770	Maize elongation f	c 640	15	1.3	2607	14	Q31985	Pisium sativum L p
568	1.3	1779	21	C41817	Arabidopsis thalia	c 641	15	1.3	2621	20	Z42002	Human endometrium
569	1.3	1798	21	Z49566	Maize MLO8 protein	c 642	15	1.3	2658	16	Q81477	Human endometrium
570	1.3	1820	21	C49561	Arabidopsis thalia	c 643	15	1.3	2668	22	F44707	RAR-alpha-403 domi
571	1.3	1827	17	C39281	Arabidopsis thalia	644	15	1.3	2673	22	F44632	Novel protein kina
572	1.3	1829	19	V44156	Arabidopsis thalia	645	15	1.3	2742	21	A12585	Novel protein kina
573	1.3	1848	21	A52771	Cytochrome p450 2C	c 646	15	1.3	2754	20	Z27471	cDNA encoding a me
574	1.3	1875	18	T91737	Human cytochrome p	c 647	15	1.3	2759	21	F21752	T. roseum DNA poly
575	1.3	1886	22	C89165	Soybean putative c	c 648	15	1.3	2763	21	Z65286	Human breast and o
576	1.3	1896	19	V59555	JC virus DNA fragm	649	15	1.3	2816	20	Z25770	Human secreted pro
577	1.3	1896	19	V59555	Human secreted pro	c 650	15	1.3	2816	21	C66029	Human psi encoding
578	1.3	1896	21	C59218	Human secreted pro	c 651	15	1.3	2841	18	V74488	Human lung cancer-
579	1.3	1901	21	C81745	Human secreted pro	c 652	15	1.3	2853	21	F21857	Staphylococcus aur
580	1.3	1903	20	Z00830	Human secreted pro	c 653	15	1.3	2876	12	Q13182	Human breast and o
581	1.3	1910	19	V59790	Human secreted pro	654	15	1.3	2927	21	A37668	T3 RNA polymerase.
582	1.3	1932	21	A60197	Hydrophobic domain	655	15	1.3	2928	13	Q29338	Human peptidase, H
583	1.3	1936	20	V74263	Human heart muscle	656	15	1.3	2928	20	V64991	RAR-alpha gene. S
584	1.3	1939	19	V52443	Streptococcus pneu	657	15	1.3	2940	10	N90124	Human RAR-alpha cD
585	1.3	1940	20	Z00850	Human secreted pro	c 658	15	1.3	2940	16	O81476	DNA of clone phRAR
586	1.3	1945	19	V59664	Human secreted pro	c 659	15	1.3	2974	20	X77838	RAR-alpha clone la
587	1.3	1987	21	C77791	Human secreted pro	c 660	15	1.3	2975	20	X77837	WMV2 strain val3 3
588	1.3	1992	21	F21852	Human cancer assoc	c 661	15	1.3	2994	19	V48269	WMV2 strain Mar6 3
589	1.3	1995	21	C51502	Human breast and o	c 662	15	1.3	3016	17	T17115	P. chrysogenum gam
590	1.3	2007	19	V59635	Arabidopsis thalia	663	15	1.3	3036	13	Q29334	Rhodospin gene. H
591	1.3	2010	21	C43217	Human secreted pro	664	15	1.3	3036	19	V20474	myl/RAR-alpha fusi
592	1.3	2017	21	A97255	Human colorectal c	c 665	15	1.3	3036	20	V64990	Human PML/RARalph
593	1.3	2021	20	Z33663	Human breast tumou	c 666	15	1.3	3089	21	A63527	Human myl/RAR-alpha
594	1.3	2048	21	A37963	Plasmid #1 DNA seq	c 667	15	1.3	3101	20	X81260	Human oligophrenin
595	1.3	2048	21	A37963	Plasmid #2 DNA seq	c 668	15	1.3	3102	21	C75441	Human ORFX ORF996
596	1.3	2052	21	A37959	Carrot promoter se	c 669	15	1.3	3119	20	X07429	Homo sapiens secre

669	1.3	3123	15	O67406	Neural alpha-caten	742	15	1.3	6146	20	X03045	Human IL-1ra BAC c
c 670	1.3	3129	14	Q43543	Rhodopsin gene. H	c 743	15	1.3	6172	10	N91825	HindIII fragment o
671	1.3	3136	21	A16631	Human secreted pro	c 744	15	1.3	6210	10	X02986	Human IL-1ra BAC c
672	1.3	3185	21	F21129	Human low adenodin	745	15	1.3	6236	18	T85728	FM22 gene associat
673	1.3	3185	21	A35007	Human adenodine re	c 746	15	1.3	6259	18	T86366	SM22 protein gene
c 674	1.3	3233	18	T79696	BRCA2 cancer susce	c 747	15	1.3	6422	20	X20576	Polynucleotide seq
c 675	1.3	3276	21	C76770	Human ORFX ORF2325	c 748	15	1.3	6536	21	C75801	Human ORFX ORF1356
676	1.3	3283	21	C59805	Human secreted pro	749	15	1.3	6590	20	X13048	Enterococcus faeca
677	1.3	3301	19	X34209	Human secreted pro	c 750	15	1.3	6727	20	X02993	Human IL-1ra BAC c
678	1.3	3308	21	A26917	Essential staphylo	751	15	1.3	6816	21	C74524	Human ORFX ORF79 p
c 679	1.3	3376	16	Q75166	AF-9 cDNA. Homo s	c 752	15	1.3	6953	17	T17116	Rhodopsin gene. H
680	1.3	3393	20	X24681	Human synaptotomal	c 753	15	1.3	7096	21	C62374	cDNA encoding a to
c 681	1.3	3427	21	C77563	Human ORFX ORF3118	c 754	15	1.3	7192	21	X93611	CAP6 gene (Bone mo
682	1.3	3511	17	T33246	PML-RAR-alpha RNA	c 755	15	1.3	7215	20	X13039	Enterococcus faeca
683	1.3	3511	17	T33259	PML-RAR-alpha DNA	c 756	15	1.3	7419	21	A49922	Human calcium chan
684	1.3	3511	21	A38656	Human PML/alpha-RA	c 757	15	1.3	7432	20	X56763	Mouse odorant bind
685	1.3	3549	20	X13796	Enterococcus faeca	c 758	15	1.3	7558	21	C87953	Candida albicans C
c 686	1.3	3621	7	N60847	Human pre-prolacti	c 759	15	1.3	7578	19	V44601	Human uncoupling p
687	1.3	3622	20	X00685	Human GPC1 DNA. H	760	15	1.3	7678	20	X12998	Enterococcus faeca
688	1.3	3703	13	Q20200	Gene encoding C-te	761	15	1.3	7720	22	C91210	Human folate recep
c 689	1.3	3744	20	X55568	Human diaphanous p	c 762	15	1.3	7918	18	T60450	Melanoma-associat
c 690	1.3	3794	21	A97363	Human colorectal c	763	15	1.3	7929	21	A58007	7929 bp Candida al
c 691	1.3	3804	18	T99544	Human haematopoiet	764	15	1.3	8365	21	F21445	Human defensin 2 p
692	1.3	3844	13	Q20199	Gene encoding C-te	765	15	1.3	8371	18	T66909	Human neuropeptide
693	1.3	3891	20	X60932	Murine cell death	766	15	1.3	8459	22	C89557	Human histone deac
694	1.3	3904	13	Q20198	Gene encoding C-te	c 767	15	1.3	8700	14	Q42541	BgIII/HpaII fragme
695	1.3	3925	12	Q11041	Encodes equine C-t	c 768	15	1.3	8700	20	X25496	S. epidermis plasm
696	1.3	3926	11	Q03736	Sequence of plasmid	769	15	1.3	8760	15	Q73473	Porcine pro-interl
697	1.3	4000	18	T91902	Mannose-1-phosphat	c 770	15	1.3	9321	21	A97904	L. mesenteroides a
c 698	1.3	4016	17	T11083	Mouse JAK3 protein	771	15	1.3	9454	21	F21444	Human defensin 2 p
c 699	1.3	4024	20	X84595	MTG16b protein cod	c 772	15	1.3	11093	18	T50855	DNA cassette for a
700	1.3	4045	13	Q20197	Gene encoding C-te	c 773	15	1.3	11628	20	X20520	Polynucleotide seq
c 701	1.3	4047	21	T38853	Human Jurkat cell	c 774	15	1.3	11715	18	T91087	Bovine lysosomal a
702	1.3	4067	21	Z38863	Human Jurkat cell	775	15	1.3	13856	18	V74342	Staphylococcus aur
703	1.3	4066	19	V10628	Human glycogen pho	776	15	1.3	14740	21	F21132	Human low adenodin
c 704	1.3	4066	21	A97913	L. mesenteroides a	777	15	1.3	14740	21	A35010	Human adenodine re
705	1.3	4146	21	C49442	Arabidopsis thalia	c 778	15	1.3	15512	19	V59470	Human ryanodin rec
c 706	1.3	4226	20	X84594	MTG16a protein cod	c 779	15	1.3	15593	19	V52185	Streptococcus pneu
707	1.3	4386	21	Z89047	Human nibrin DNA.	c 780	15	1.3	16798	20	X13132	Enterococcus faeca
708	1.3	4386	21	Z34997	NBS1 gene associat	781	15	1.3	19702	19	V52140	Streptococcus pneu
709	1.3	4415	14	Q46293	Phenylalanine ammo	c 782	15	1.3	20387	19	V62159	HSV-2 strain SB5 C
710	1.3	4415	21	C61194	Pea PSPAL1 DNA seq	783	15	1.3	20710	17	T30681	Kaposi's sarcoma a
711	1.3	4440	18	T89487	Human A20 protein	784	15	1.3	20710	17	T16806	Kaposi's sarcoma a
c 712	1.3	4444	20	X20635	Polynucleotide seq	c 785	15	1.3	26281	21	Z60929	Nucleotide sequenc
713	1.3	4621	18	T74034	Mouse male enhance	c 786	15	1.3	26338	19	V62134	Polynucleotide seq
714	1.3	4621	18	X04132	Mouse male enhance	787	15	1.3	32768	20	X20515	Polynucleotide seq
715	1.3	4655	21	A34784	Human adenodine re	788	15	1.3	35100	20	V73803	KSHV LTR DNA (nucl
c 716	1.3	4718	18	V74610	Staphylococcus aur	c 789	15	1.3	43069	21	Z36335	Genomic sequence o
717	1.3	4751	20	Z09496	Human heart tissue	790	15	1.3	49272	19	V35000	Mycobacteriophage
718	1.3	4783	21	C76742	Human ORFX ORF2297	c 791	15	1.3	50000	21	A96367	Polymorphic repeat
c 719	1.3	4798	21	C74503	Human ORFX ORF58 p	792	15	1.3	50000	21	A96367	Polymorphic repeat
720	1.3	4811	20	X08942	GATA-2 promoter re	c 793	15	1.3	65632	21	A81502	N. meningitidis pa
c 721	1.3	4837	21	C77142	Human ORFX ORF2697	794	15	1.3	65921	21	Z89046	Human nibrin DNA.
c 722	1.3	4911	21	C48475	Arabidopsis thalia	c 795	15	1.3	65921	21	Z89046	Human nibrin DNA.
723	1.3	5053	19	V22816	Nucleotide sequenc	c 796	15	1.3	85680	21	F22299	BAC containing rep
724	1.3	5055	19	V17236	DNA from a region	c 797	15	1.3	96988	21	F22299	BAC containing rep
725	1.3	5057	21	C83734	Mouse laminin 5 CD	c 798	15	1.3	99960	21	Z50905	Human TBC-1 partia
c 726	1.3	5117	21	A49920	Human calcium chan	799	15	1.3	116277	20	X20249	Borrelia burgdorfe
727	1.3	5159	21	C83733	Mouse laminin 5 CD	c 800	15	1.3	116624	19	V52850	Human eyal gene co
728	1.3	5175	20	X40200	Restin encoding ge	c 801	15	1.3	117213	19	V62176	HSV-2 strain SB5 C
c 729	1.3	5194	20	X25885	C.albicans alpha-I	802	15	1.3	129021	21	F22296	BAC containing rep
c 730	1.3	5333	21	A57410	DNA sequence which	803	15	1.3	134525	11	Q04525	Total base sequenc
731	1.3	5734	21	A57958	5734 bp Candida al	804	15	1.3	137507	19	V19941	KSHV long unique c
c 732	1.3	5778	5	N40033	Sequence of portio	c 805	15	1.3	143068	21	F21105	Human low adenodin
c 733	1.3	5822	20	V62933	Human mdia Rho tar	c 806	15	1.3	143068	21	F21272	Human low adenodin
734	1.3	5822	20	V62934	Human mdia Rho tar	c 807	15	1.3	143068	21	A34983	Human adenodine re
c 735	1.3	5828	20	X84592	AML1-MTG16 fusion	c 808	15	1.3	143068	21	A35150	Human adenodine re
736	1.3	5858	12	Q10378	Encodes human 160k	c 809	15	1.3	149412	21	A35150	Human adenodine re
737	1.3	6000	16	Q86478	Human PTP-OB. Hom	c 810	15	1.3	152740	21	F21273	Human low adenodin
738	1.3	6000	16	T85389	Human protein tyro	c 811	15	1.3	209273	21	F21437	Human factor-relat
739	1.3	6000	20	X06095	Human protein tyro	812	15	1.3	237326	19	V57903	Hereditary haemoch
c 740	1.3	6035	18	V45883	Staphylococcus aur	813	15	1.3	273254	21	C81914	Chlamydia pneumoni
c 741	1.3	6056	20	X84593	AML1-MTG16 fusion	c 814	15	1.3	273254	21	C81914	Chlamydia pneumoni

815 15 1.3 349980 21 F21544  
 c 816 15 1.3 611590 21 F22303  
 817 15 1.3 910715 20 X20248  
 c 818 15 1.3 910715 20 X20248  
 819 15 1.3 1230025 20 X31990  
 c 820 15 1.3 1230025 20 X31990  
 821 15 1.3 1437668 21 A81490  
 822 15 1.3 1664976 19 V21209  
 c 823 15 1.3 1830121 17 T42063

## ALIGNMENTS

## RESULT 1

ID V15418 standard; cDNA; 1119 BP.

AC V15418;

XX 11-JUN-1998 (first entry)

DE Human dendritic cell chemokine receptor encoding cDNA.

XX Human; thymus expressed chemokine; TECK; MIP-3alpha; MIP-3beta;  
 KW receptor; dendritic cell; macrophage; inflammation; asthma; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..1098

FT /\*tag= a

FT /product= "DC CR"

FT /note= "dendritic cell chemokine receptor"

FT 577..579

FT /\*tag= b

FT /note= "encodes His or Gln, but is stated as Gln in

the protein (shown in W48086)"

PN WO9801557-A2.

PD 15-JAN-1998.

XX 02-JUL-1997; 97WO-US10819.

XX 04-JUN-1997; 97US-0048593.

PR 05-JUL-1996; 96US-0675814.

PR 11-OCT-1996; 96US-0028329.

XX (SCHE ) SCHERING CORP.

XX Gish KC, Schall TJ, Vicari A, Wang W, Zlotnik A;

XX WPI; 1998-101054/09.

DR P-PSDB; W48086.

PT Novel chemokines, e.g. thymus expressed chemokine - used for  
 PT treating inflammatory conditions including asthma.

XX Claim 5; Page 92-94; 202pp; English.

XX The present sequence encodes a human dendritic cell chemokine receptor.  
 CC Antibodies which bind to the protein can be used in detecting or  
 CC diagnosing various immunological conditions related to expression  
 CC of the protein. The nucleic acid can be used for screening and  
 CC isolating DNA clones for the chemokines, especially from other  
 CC species. The chemokine can be used in the treatment of conditions  
 CC associated with abnormal physiology or development, including  
 CC inflammatory conditions such as asthma.

XX Sequence 1119 BP; 261 A; 268 C; 257 G; 332 T; 1 other;

Query Match 95.4%; Score 1068; DB 19; Length 1119;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1118; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTTTTCGACTCCAGTGAAGATTATTTGTGTGCTCAATATCTTCAATCTAGTTG 60  
 Db 1 atgttttcgactccagtgagattattttgtgcagtcataacttcatattactcagtg 60  
 Qy 61 ATCTCAGATGTTACTGTCTCTTCCAGAGGTCCAGGAGTCTCCAGGCTATTGTGAC 120  
 Db 61 attctgagatgttactgtctctcttcagaggagtcaggagtcctccaggtatttgtac 120  
 Qy 121 CGAATTGCCTACTCTGTGATCTGTGCTTCTGGGCTCTCTGGGGAATATCTGTGGTGATC 180  
 Db 121 cgaattgcctactctctgtatctgttcttgcctcctctggggaattctctgtgtgagtc 180  
 Qy 181 ACCTTTGCTTTTATAAGAGGCCAGGTCTATGACAGACGTCTATCTCTTGAACATGGCC 240  
 Db 181 acctttgcttttataagaagccaggtctatgacagacgtctctctcttgaacatggcc 240  
 Qy 241 ATTCGAGACATCCTCTTTGTTCTTACTCTCCCATTTCTGGCAGTGAGTCATGCCACTGGT 300  
 Db 241 attgcagacatcctctttgttcttactctcccatctctggcagtgagtcagtcagtcaggt 300  
 Qy 301 GCGTGGGTTTTCAGCAATGCCACGTGCAAGTTGCTAAAAGGCATCTATGCCATCAACTTTT 360  
 Db 301 gcgtgggttttcagcaatgccacgtgcaagttgtctaaaaggcatctatgccatcaacttt 360  
 Qy 361 AACTCGGGGATGCTGCTCCTGACTTGCATTAGCATGGACGGGTACATGCGCATGTACAG 420  
 Db 361 aactcggggagtgtctctctgactgtcattagcatggacgggtacatgcgcatgtacag 420  
 Qy 421 CGGACTAAGTCATTCGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCCTT 480  
 Db 421 gcgactaagtcattccgctccgagtcacgaacactaccgcgcagcaaaatcatctgcctt 480  
 Qy 481 GTGTGTGGGGCTGTTCAGTTCATCTCAGTCACTCACTCAACTTTTCTCTCAACCAAAATAC 540  
 Db 481 gttgtgtggggctgtcagtcacatcatctccagctcaactttgtcttcaacaaaatac 540  
 Qy 541 AACACCCCAAGCAGCATGCTGTGTAACCCAGTACCAAACTGTCTCGGAGCCCATCAGG 600  
 Db 541 aacaccccaagcagcagtgctgtgtaacccaagtaccanactgtctcggagccatcagg 600  
 Qy 601 TGGAGCTGCTGATGTGGGCTTGAGCTACTCTTTGGTTCTTTATCCCTTTGATGTTTC 660  
 Db 601 tggagctgctgatgttgggcttgagctactctttggtttcttcccttttgatgttc 660  
 Qy 661 ATGATATTTTGTACAGGTTTCATTTGTCAAAACCTTGGTGCAGCTCAGAAATCTTAAAGG 720  
 Db 661 atgatatTTTGTACAGGTTTCATTTGTCAAAACCTTGGTGCAGCTCAGAAATCTTAAAGG 720  
 Qy 721 CACAAAGCCATCCGTGTAATCATAGCTGTGGTCTTGTGTTTCTGGCTTGTGAGATTCCT 780  
 Db 721 cacaaagccatccgtgtaatcatagctgtggctgtgtgtgttctgttctgttcagattcct 780  
 Qy 781 CATACATGGTCTCTGCTTGTGAGCGCTGCTAATTTGGGTAAATGAACCATCTTGGCCAG 840  
 Db 781 catacatggctctgtgtgagcgtgtgtaatttgggttaaaatgaacccatcctgcag 840  
 Qy 841 AGCGAAAAGCTAATTTGGCTATACGAAACTGTTCACAGAGTCTCTGGCTTTCTTGCCTGTC 900  
 Db 841 agcgaagctaat199ctatacgaataactgtccaggaagtcctcttcttccctgcactgc 900  
 Qy 901 TGCTTGAACCCCTGTGCTCTACGCTTTTATTGGGAGAAAGTTTCAAGAACTACTTCTTGAAG 960  
 Db 901 tgcttgaacccctgtgctctacgcttttatttgggcagaagtcagaactactttctgaag 960  
 Qy 961 ATCTTGAAGGACCTGTGGTGTGTGAGAAGAGTCAAGTCTCAGGCTTCTCTGTGCCC 1020  
 Db 961 atcttgaaggacctgtgtgtgagaaggaagtaacagtcctcaggtcttctcctgtgcc 1020  
 Qy 1021 GGGAGGTACTCAGAAAACATTTCTCGGCAGACCAGTGAGACCCGAGATACGACAAATGCG 1080

```
Db 1021 gggaggtaactagaaaaacattctcggcagaccagtgagaccgcagataaacgacaatgcg 1080
|||||
QY 1081 TCGTCTTCACATATGATAGAGAAAGCTGAGTCTCCCTAA 1119
|||||
Db 1081 tgcgtctcactatgtagagaaagctgagtcctccctaa 1119
|||||

RESULT 2
A51971
ID A51971 standard; cDNA; 1098 BP.
XX
AC A51971;
XX
DT 04-DEC-2000 (first entry)
XX
DE Primate (human) chemokine receptor CCR6 coding sequence.
XX
KW MIP-3-alpha; chemokine; CCR6 receptor; ligand; cytostatic; agonist;
KW antagonist; anti-psoriatic; dermatological; immunosuppressive;
XX anti-inflammatory; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1098
FT FT /*tag= a
FT FT /transl_except= (pos:577..579, aa:Xaa)
FT FT /note= "Xaa is not defined"
FT FT /product= CCR6
XX
PN W0200046248-A1.
XX
XX 10-AUG-2000.
XX
XX 02-FEB-2000; 2000WO-US00511.
XX
XX 03-FEB-1999; 99US-0244281.
XX
XX (SCHE ) SCHERING CORP.
XX
XX Oldham ER, Honey B, Dieu-Nosjean M, Caux C, Zlotnik A;
XX
XX WPI; 2000-543477/49.
XX
XX P-PSDB; Y97077.
XX
XX Novel methods for modulating the migration of cells within or to the
XX skin using MIP-3-alpha or CCR6 agonists or antagonists, useful for
XX treating skin disorders, e.g. cancer
XX
XX Disclosure; Page 51-53; 61pp; English.
XX
XX The MIP-3-alpha chemokine is expressed in inflamed skin cells. This
XX chemokine is the ligand for the CCR6 receptor. MIP-3-alpha or CCR6
XX agonists or antagonists can be used to modulate the migration of a cell
XX within or to the skin of a mammal. MIP-3-alpha can be used to purify a
XX population of cells expressing a MIP-3-alpha receptor. The methods are
XX useful for treating a mammalian subject with a skin disease or condition,
XX e.g. cancer, metastasis, psoriasis, atopic or contact dermatitis,
XX systemic lupus erythematosus and lichen ruber planus, or for treating
XX skin transplants or grafts.
XX
XX Sequence 1098 BP; 254 A; 263 C; 253 G; 327 T; 1 other;

Query Match 93.6%; Score 1047; DB 21; Length 1098;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1097; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTTTCAGCTCCAGTGAAGATATTTTGTGTCAGTCAATATCTCATATTAATCTAGTTG 60
|||||
Db 1 atgttttcagctccagtgagattatttgtgtcagtcacatacttcattactcagttg 60
```

RESULT 3  
 ID A47508/c  
 XX A47508 standard; cDNA; 24 BP.  
 AC  
 XX A47508;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Primer for amplifying CCR6 chemokine.  
 XX  
 KW Secondary lymphoid chemokine; SLC; cancer;  
 KW hyperproliferative disorders; prostatic hyperplasia;  
 KW proliferative breast disease; proliferative retinopathy;  
 KW melanoma; breast cancer; cancer; metastases; suppression;  
 KW angiogenesis; tumorigenesis; inflammation; immune response;  
 KW chemotaxis; graft rejection; autoimmune disease; primer; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200038706-A2.  
 XX  
 PD 06-JUL-2000.  
 XX  
 PF 28-DEC-1999; 99WO-US31096.  
 XX  
 PR 31-DEC-1998; 98US-0114498.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Keting C, Xin H, Chan VWF, Kothakota S, Williams LT, Winter JA;  
 XX  
 DR WPI; 2000-465631/40.  
 XX  
 PT Treating cancer or hyperproliferative disorder and modulating dendritic  
 PT cell function in a mammal involves administering secondary lymphoid  
 PT chemokine to the mammal  
 XX  
 PS Disclosure; Page 28; 53pp; English.  
 XX  
 CC Secondary lymphoid chemokines (SLC's), variants, fragments, and the  
 CC polynucleotides encoding the chemokines, variants and fragments,  
 CC anti-SLC antibodies or ligands for the CCR7 receptor can be used to  
 CC modulate dendritic cell function in a mammal which results in a  
 CC decreased primary immune response. SLC can be used to treat cancer or  
 CC hyperproliferative disorders such as prostatic hyperplasia.  
 CC proliferative breast diseases, proliferative retinopathy or pigmented  
 CC skin lesions. SLC is also useful for treating solid tumours such as  
 CC melanoma, breast cancer, tumours of the head and neck, cancers or  
 CC metastases of ovary, endometrium, urinary tract, stomach, testicle,  
 CC prostate, lung, bladder, pancreas, bone, liver, colon or rectum, or  
 CC metastases of unknown primary origin. SLC can also be used to  
 CC suppress angiogenesis particularly angiogenesis involved in cancer,  
 CC tumorigenesis, metastases and tumour growth, and for mediating  
 CC recruitment of leukocytes into sites of inflammation and immune  
 CC responses, particularly, the chemotaxis of dendritic and other cells.  
 CC SLC is also useful in preventing graft rejection, prevention and  
 CC treatment of the autoimmune diseases and for enhancing an immune  
 CC response. Two primers (A47507, A47508) were used to amplify the  
 CC sequence encoding the chemokine CCR6.  
 XX  
 SQ Sequence 24 BP; 3 A; 7 C; 5 G; 9 T; 0 other;  
 XX  
 Query Match 2.1%; Score 24; DB 21; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 0.018;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1055 GTGAGACCCGACATACACCAATG 1078  
 DB 24 GTGAGACCCGACATACACCAATG 1  
 XX  
 RESULT 4

V31479/c  
 ID V31479 standard; DNA; 34 BP.  
 XX  
 AC V31479;  
 XX  
 DT 07-SEP-1998 (first entry)  
 XX  
 DE PCR primer HSU45984-Not1R used in the course of the invention.  
 XX  
 KW Human; liver and activation regulated chemokine; LARC; treatment;  
 KW diagnosis; disease; inflammation; immune reaction; cancer; AIDS;  
 KW autoimmune disease; LARC binding; CC chemokine receptor; GPR-CY4;  
 KW PCR primer; ss.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9817800-A1.  
 XX  
 PD 30-APR-1998.  
 XX  
 PF 24-JUL-1997; 97WO-JP02557.  
 XX  
 PR 31-JAN-1997; 97JP-0018038.  
 PR 17-OCT-1996; 96JP-0274998.  
 XX  
 PA (SHIO ) SHIONOGI & CO LTD.  
 XX  
 PI Imai T, Nomiya H, Yoshie O;  
 XX  
 DR WPI; 1998-261496/23.  
 XX  
 PT New human liver and activation chemokine - useful for, e.g.  
 PT treatment and diagnosis of inflammatory and immunological disorders  
 XX  
 PS Example 4; Page 24; 52pp; Japanese.  
 XX  
 CC PCR primers V31478-79 were used in the course of the invention. The  
 CC specification describes a human liver and activation regulated  
 CC chemokine (LARC). LARC can be used for treating and diagnosing diseases  
 CC involving inflammation and immune reactions, including hereditary  
 CC diseases, cancer, AIDS and autoimmune diseases. LARC can also be used in  
 CC screening for potential agonists, inverse agonists and antagonists of  
 CC LARC by measuring their effect on the binding reaction of LARC with  
 CC CC chemokine receptors such as GPR-CY4.  
 XX  
 SQ Sequence 34 BP; 8 A; 11 C; 11 G; 4 T; 0 other;  
 XX  
 Query Match 2.1%; Score 23; DB 19; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 0.059;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1076 ATGCGTCGTCCTTCACATATGTA 1098  
 DB 34 ATGCGTCGTCCTTCACATATGTA 12  
 XX  
 RESULT 5  
 Z60074/c  
 ID Z60074 standard; DNA; 21 BP.  
 XX  
 AC Z60074;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Reverse PCR primer -1080/CCR6 used to amplify chemokine receptor CCR6.  
 XX  
 KW Chemokine receptor; CCR6; PCR primer; dendritic cell; disease treatment;  
 KW immune response initiation; infection; cancer; autoimmune disease;  
 KW tissue rejection; allergy; ss.  
 XX  
 OS Homo sapiens.

XX EP974357-A1.  
 XX  
 XX PD 26-JAN-2000.  
 XX  
 XX PF 16-JUL-1998; 98EP-0401799.  
 XX  
 XX PR 16-JUL-1998; 98EP-0401799.  
 XX  
 XX PA (SCHE ) SCHERING-PLOUGH.  
 XX  
 XX PI Caux C, Vanbervliet B, Lebecque S, Vicari A, Dieu M;  
 XX  
 XX WPI; 2000-118300/11.  
 XX  
 XX Use of chemokines capable of directing migration of dendritic cells,  
 XX PT useful for treating microbial infections, cancer and autoimmune  
 XX PT diseases -  
 XX  
 XX PS Disclosure; Column 12; 16pp; English.  
 XX  
 XX This sequence represents a PCR primer used to amplify the chemokine  
 XX receptor CCR6 coding sequence. The PCR product is used in the analysis of  
 XX dendritic cell response to different chemokines. CCR6 is a receptor for  
 XX the macrophage inflammation protein 3 alpha (MIP 3alpha). The invention  
 XX relates to the use of chemokines which are capable of directing dendritic  
 XX cells, in the manufacture of a medicament for the treatment of a disease  
 XX state. Methods are included for treating diseases by facilitating or  
 XX inhibiting the migration or activation of antigen-presenting dendritic  
 XX cells. The chemokines can be used to initiate, amplify or modulate an  
 XX immune response. The chemokines are useful for the treatment of disease  
 XX states e.g. a bacterial, viral, fungal or parasitic infection, cancer  
 XX (especially melanoma, breast, pancreatic, colon, lung, glioma,  
 XX hepatocellular, endometrial, gastric, intestinal, renal, prostate,  
 XX thyroid, ovarian, testicular, liver, head and neck, colorectal,  
 XX oesophagus, stomach, eye, bladder, glioblastoma and metastatic  
 XX carcinomas), autoimmune disease, tissue rejection or an allergy.  
 XX  
 XX Sequence 21 BP; 6 A; 3 C; 8 G; 4 T; 0 other;  
 XX

Query Match 1.9%; Score 21; DB 21; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.63;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 994 TACAAGTCTCTCAGGCTTCTCC 1014  
 |||||  
 Db 21 TACAAGTCTCTCAGGCTTCTCC 1

RESULT 6  
 ID 260077/c  
 XX ID 260077 standard; DNA; 21 BP.  
 XX  
 XX AC 260077;  
 XX  
 XX DT 25-APR-2000 (first entry)  
 XX  
 XX DE Probe -622/CCR6 for chemokine receptor CCR6.  
 XX  
 XX KW Chemokine receptor; CCR6; probe; dendritic cell; disease treatment;  
 XX immune response initiation; infection; cancer; autoimmune disease;  
 XX tissue rejection; allergy; ss.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX PN EP974357-A1.  
 XX  
 XX PD 26-JAN-2000.  
 XX  
 XX PF 16-JUL-1998; 98EP-0401799.  
 XX  
 XX PR 16-JUL-1998; 98EP-0401799.  
 XX  
 XX PI 16-JUL-1998; 98EP-0401799.

XX (SCHE ) SCHERING-PLOUGH.  
 XX  
 XX Caux C, Vanbervliet B, Lebecque S, Vicari A, Dieu M;  
 XX  
 XX WPI; 2000-118300/11.  
 XX  
 XX Use of chemokines capable of directing migration of dendritic cells,  
 XX PT useful for treating microbial infections, cancer and autoimmune  
 XX PT diseases -  
 XX  
 XX PS Disclosure; Column 12; 16pp; English.  
 XX  
 XX This sequence represents a probe for the chemokine receptor CCR6. The  
 XX probe is used in the analysis of dendritic cell response to different  
 XX chemokines. CCR6 is a receptor for the macrophage inflammation protein 3  
 XX alpha (MIP 3alpha). The invention relates to the use of chemokines which  
 XX are capable of directing dendritic cells, in the manufacture of a  
 XX medicament for the treatment of a disease state. Methods are included for  
 XX treating diseases by facilitating or inhibiting the migration or  
 XX activation of antigen-presenting dendritic cells. The chemokines can be  
 XX used to initiate, amplify or modulate an immune response. The chemokines  
 XX are useful for the treatment of disease states e.g. a bacterial, viral,  
 XX fungal or parasitic infection, cancer (especially melanoma, breast,  
 XX pancreatic, colon, lung, glioma, hepatocellular, endometrial, gastric,  
 XX intestinal, renal, prostate, thyroid, ovarian, testicular, liver, head  
 XX and neck, colorectal, oesophagus, stomach, eye, bladder, glioblastoma and  
 XX metastatic carcinomas), autoimmune disease, tissue rejection or an  
 XX allergy.  
 XX  
 XX Sequence 21 BP; 1 A; 3 C; 7 G; 10 T; 0 other;  
 XX

Query Match 1.9%; Score 21; DB 21; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.63;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 535 AAATACACACCCCAAGGCAGC 555  
 |||||  
 Db 21 AAATACACACCCCAAGGCAGC 1

RESULT 7  
 A47507  
 ID A47507 standard; cDNA; 20 BP.  
 XX  
 XX AC A47507;  
 XX  
 XX DT 20-OCT-2000 (first entry)  
 XX  
 XX DE Primer for amplifying CCR6 chemokine.  
 XX  
 XX KW Secondary lymphoid chemokine; SLC; cancer;  
 XX hyperproliferative disorders; prostatic hyperplasia;  
 XX proliferative breast disease; proliferative retinopathy;  
 XX melanoma; breast cancer; cancer; metastases; suppression;  
 XX angiogenesis; tumorigenesis; inflammation; immune response;  
 XX chemotaxis; graft rejection; autoimmune disease; primer; ss.  
 XX  
 XX OS Synthetic.  
 XX  
 XX PN WO200038706-A2.  
 XX  
 XX PD 06-JUL-2000.  
 XX  
 XX PF 28-DEC-1999; 99WO-US31096.  
 XX  
 XX PR 31-DEC-1998; 98US-0114498.  
 XX  
 XX PA (CHIR ) CHIRON CORP.  
 XX  
 XX PI Keting C, Xin H, Chan VWF, Kothakota S, Williams LT, Winter JA;  
 XX

DR WPI; 2000-465631/40.  
 XX Treating cancer or hyperproliferative disorder and modulating dendritic  
 PT cell function in a mammal involves administering secondary lymphoid  
 PT chemokine to the mammal.  
 XX  
 XX Disclosure; Page 28; 53pp; English.  
 XX  
 CC Secondary lymphoid chemokines (SLC's), variants, fragments, and the  
 CC polynucleotides encoding the chemokines, variants and fragments,  
 CC anti-SLC antibodies or ligands for the CCR7 receptor can be used to  
 CC modulate dendritic cell function in a mammal which results in a  
 CC decreased primary immune response. SLC can be used to treat cancer or  
 CC hyperproliferative disorders such as prostatic hyperplasia,  
 CC proliferative breast diseases, proliferative retinopathy or pigmented  
 CC skin lesions. SLC is also useful for treating solid tumours such as  
 CC melanoma, breast cancer, tumours of the head and neck, cancers or  
 CC metastases of ovary, endometrium, urinary tract, stomach, testicle,  
 CC prostate, lung, bladder, pancreas, bone, liver, colon or rectum, or  
 CC metastases of unknown primary origin. SLC can also be used to  
 CC suppress angiogenesis particularly angiogenesis involved in cancer,  
 CC tumourigenesis, metastases and tumour growth, and for mediating  
 CC recruitment of leukocytes into sites of inflammation and immune  
 CC responses, particularly, the chemotaxis of dendritic and other cells.  
 CC SLC is also useful in preventing graft rejection, prevention and  
 CC treatment of the autoimmune diseases and for enhancing an immune  
 CC response. Two primers (A47507, A47508) were used to amplify the  
 CC sequence encoding the chemokine CCR6.  
 XX  
 SQ Sequence 20 BP; 7 A; 6 C; 5 G; 2 T; 0 other;

Query Match 1.8%; Score 20; DB 21; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.1; Indels 0; Gaps 0;  
 Matches 20; Conservative 0; Mismatches 0;

Qy 831 ATCTGCGCAGAGCGAAAGC 850  
 |||||  
 Db 1 atctgcgagcgaaagc 20

RESULT 8  
 Z00369/C  
 ID Z00369 standard; DNA; 2277 BP.  
 XX  
 AC Z00369;  
 XX  
 DT 26-OCT-1999 (first entry)  
 XX  
 DE Nucleotide sequence of murine mmFATP5.

XX Fatty acid transport protein; FATP; long chain fatty acid; LCFA; murine;  
 KW fatty acid; FATP biosynthesis; obesity; diabetes; heart disease; ss.  
 XX  
 XX Mus sp.

OS  
 XX WO9936537-A2.  
 XX  
 PD 22-JUL-1999.  
 XX  
 XX 14-JAN-1999; 99WO-US00182.  
 XX  
 PR 14-JAN-1999; 99US-0232201.  
 PR 15-JAN-1998; 98US-0071374.  
 PR 20-JUL-1998; 98US-0093491.  
 PR 04-DEC-1998; 98US-0110941.  
 PR 14-JAN-1999; 99US-0232195.  
 PR 14-JAN-1999; 99US-0232197.  
 PR 14-JAN-1999; 99US-0232200.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA (WHEED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX

PI Gimeno RE, Hirsch DJ, Lodish HF, Stahl A, Tartaglia LA;  
 XX WPI; 1999-444398/37.  
 DR P-PSDB; Y14959.  
 XX  
 XX Fatty acid transport proteins and related polynucleotides, useful  
 PT for treating obesity, diabetes and heart disease  
 XX  
 XX Example 1; Fig 70; 255pp; English.

XX The invention provides a family of fatty acid transport proteins (FATPs)  
 CC that mediate transport of long chain fatty acids (LCFAs) across cell  
 CC membranes into cells. Human and murine FATP proteins and nucleic acids  
 CC encoding the proteins are provided. The FATP proteins can be produced  
 CC by standard recombinant methodology. Fatty acid uptake by cells can be  
 CC modulated by modulating biosynthesis of FATP proteins especially FATP6.  
 CC In particular, antisense oligonucleotides can be used to modulate FATP  
 CC biosynthesis. Modulation of FATP6 is useful for inhibiting fatty acid  
 CC uptake in cardiac muscle of humans. Agents can be directed to cardiac  
 CC muscle or liver by administration of a complex of the agent and a FATP6  
 CC binding moiety. DNA encoding FATP proteins can be used as a reference  
 CC used in detecting variant alleles or homologues. Altering the LCFA uptake  
 CC by administering an inhibitor or enhancer of FATP transport function in  
 CC the small intestine can decrease or increase calories available as fats,  
 CC and can decrease or increase circulating fatty acids. Blocking the  
 CC function of FATP4 and also FATP2, is useful for treating obesity,  
 CC diabetes and heart disease.

XX Sequence 2277 BP; 521 A; 593 C; 609 G; 554 T; 0 other;

Query Match 1.7%; Score 19; DB 20; Length 2277;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 876 AGAAGTCCTGCTTCCTG 894  
 |||||  
 Db 1540 AGAAGTCCTGCTTCCTG 1522

RESULT 9  
 Z00345/C  
 ID Z00345 standard; DNA; 2277 BP.

XX  
 AC Z00345;  
 XX  
 DT 26-OCT-1999 (first entry)  
 XX  
 DE Nucleotide sequence of murine mmFATP5.

XX Fatty acid transport protein; FATP; long chain fatty acid; LCFA; murine;  
 KW fatty acid; FATP biosynthesis; obesity; diabetes; heart disease; ss.  
 XX  
 XX Mus musculus.

OS  
 XX WO9936537-A2.  
 XX  
 PD 22-JUL-1999.  
 XX  
 XX 14-JAN-1999; 99WO-US00182.  
 XX  
 PR 14-JAN-1999; 99US-0232201.  
 PR 15-JAN-1998; 98US-0071374.  
 PR 20-JUL-1998; 98US-0093491.  
 PR 04-DEC-1998; 98US-0110941.  
 PR 14-JAN-1999; 99US-0232195.  
 PR 14-JAN-1999; 99US-0232197.  
 PR 14-JAN-1999; 99US-0232200.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA (WHEED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX  
 XX Gimeno RE, Hirsch DJ, Lodish HF, Stahl A, Tartaglia LA;

XX WPI: 1999-444398/37.  
 DR P-PSDB; Y14935.  
 XX Fatty acid transport proteins and related polynucleotides, useful  
 PT for treating obesity, diabetes and heart disease  
 XX Example 1; Fig 12; 255pp; English.  
 PS  
 XX The invention provides a family of fatty acid transport proteins (FATPs)  
 CC that mediate transport of long chain fatty acids (LCFAs) across cell  
 CC membranes into cells. Human and murine FATP proteins and nucleic acids  
 CC encoding the proteins are provided. The FATP proteins can be produced  
 CC by standard recombinant methodology. Fatty acid uptake by cells can be  
 CC modulated by modulating biosynthesis of FATP proteins especially FATP6.  
 CC In particular, antisense oligonucleotides can be used to modulate FATP  
 CC biosynthesis. Modulation of FATP6 is useful for inhibiting fatty acid  
 CC uptake in cardiac muscle of humans. Agents can be directed to cardiac  
 CC muscle or liver by administration of a complex of the agent and a FATP6  
 CC binding moiety. DNA encoding FATP proteins can be used as a reference  
 CC used in detecting variant alleles or homologues. Altering the LCFA uptake  
 CC by administering an inhibitor or enhancer of FATP transport function in  
 CC the small intestine can decrease or increase calories available as fats,  
 CC and can decrease or increase circulating fatty acids. Blocking the  
 CC function of FATP4 and also FATP2, is useful for treating obesity,  
 CC diabetes and heart disease.  
 XX Sequence 2277 BP; 521 A; 592 C; 610 G; 554 T; 0 other;

Query Match 1.7%; Score 19; DB 20; Length 2277;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 876 AGAAGTCCTGGCTTCCCTG 894  
 |||||  
 Db 1540 AGAAGTCCTGGCTTCCCTG 1522

RESULT 10  
 Q05753  
 ID Q05753 standard; DNA; 3293 BP.

XX Key

AC Q05753;

XX 02-JAN-1991 (first entry)

XX Microspore-specific clone L16 from Brassica napus.

XX microspore; controllable male sterility; blocking pollen formation;  
 KW anti-sense DNA; ss;

OS Brassica napus spp. oleifera Westar.

XX Key Location/Qualifiers

FT exon 1..124

FT /tag= a

FT /number=1

FT /note="homologous to exon #1 of clone L10."

FT intron 125..689

FT /tag= b

FT /number=1

FT 690..1796

FT /tag= c

FT /number=2

FT /note="homologous to exon #2 of clone L10"

FT intron 1797..1911

FT /tag= d

FT /number=2

FT 1912..2212

FT /tag= e

FT /number=3

PN W09008828-A.  
 XX  
 PD 09-AUG-1990.  
 XX  
 PF 02-FEB-1990; 90WO-CA00037.  
 XX  
 PR 03-FEB-1989; 89EP-0301053.  
 PR 02-FEB-1989; 89NZ-0227835.  
 PR 03-FEB-1989; 89US-0306438.  
 XX  
 PA (PALA-) PALADIN HYBRIDS INC.  
 XX  
 PI Fabijanski S, Albani D, Robert LS, Arnison PG;  
 XX  
 DR WPI: 1990-260937/34.  
 DR P-PSDB; R06519.  
 XX  
 XX Hybrid seed prodn. from plants with controllable male sterility -  
 PT induced by insertion of anti-sense DNA which blocks formation of  
 PT pollen.  
 XX  
 PS Disclosure; Fig 3c; 207pp; English.  
 XX  
 CC Clone L16 was identified as microspore-specific and critical to  
 CC pollen formation and function. The sequence does not have a  
 CC promoter sequence and so is only used as a source of coding  
 CC sequences for anti-sense RNA constructs. It may be cloned  
 CC into a plasmid as a pollen-specific gene in the anti-sense  
 CC orientation with respect to the promoter. The construct is  
 CC used to transform Brassica sp.  
 CC In addition the plasmid carries a resistance gene to a chemical  
 CC agent or stress. The sense and anti-sense genes are regulated such  
 CC that they are expressed at about the same time as each other.  
 CC Transfected plants are pollinated and selected according to  
 CC presence of the resistance gene. The promoter is inducible so that  
 CC plants are male-sterile only in the presence of the appropriate  
 CC inducer.  
 CC See also Q05749, Q05752 and Q05754-Q05758.  
 XX  
 SQ Sequence 3293 BP; 1031 A; 701 C; 579 G; 969 T; 13 other;

Query Match 1.7%; Score 19; DB 11; Length 3293;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1097 GATAGAAAGCTGAGTCCTCC 1115  
 |||||

Db 2811 gatagaaagctgagtcctcc 2829

RESULT 11

T48862

ID T48862 standard; DNA; 3293 BP.

XX

AC T48862;

XX 07-APR-1997 (first entry)

XX Brassica napus microspore-specific clone L16 fragment.

XX Male sterile plant; hybrid seed; pollen; microspore; oilseed;  
 KW canola; rape; antisense RNA; stress resistance;  
 KW herbicide resistance; transgenic plant; crop protection; ds.

OS Brassica napus.

XX Key Location/Qualifiers

FT exon 1..124

FT /tag= a

FT intron 125..688

FT /tag= b

FT exon 689..1793

FT intron /\*tag= c  
 FT 1794...1909 /\*tag= d  
 FT exon 1910...2212 /\*tag= e  
 FT misc\_difference 14...15 /\*tag= f  
 FT /\*note= "nn bases at positions 14-15 are not identified in the specification"  
 FT misc\_difference 891 /\*tag= g  
 FT /\*note= "base n at position 891 is not identified in the specification"  
 FT misc\_difference 1578 /\*tag= h  
 FT /\*note= "base n at position 1578 is not identified in the specification"  
 FT misc\_difference 1608 /\*tag= i  
 FT /\*note= "base n at position 1608 is not identified in the specification"  
 FT misc\_difference 1837 /\*tag= j  
 FT /\*note= "base n at position 1837 is not identified in the specification"  
 FT misc\_difference 1840 /\*tag= k  
 FT /\*note= "base n at position 1840 is not identified in the specification"  
 FT misc\_difference 1993 /\*tag= l  
 FT /\*note= "base n at position 1993 is not identified in the specification"  
 FT misc\_difference 2613 /\*tag= m  
 FT /\*note= "base n at position 2613 is not identified in the specification"  
 FT misc\_difference 2646 /\*tag= n  
 FT /\*note= "base n at position 2646 is not identified in the specification"  
 FT misc\_difference 2996 /\*tag= o  
 FT /\*note= "base n at position 2996 is not identified in the specification"  
 FT misc\_difference 3057 /\*tag= p  
 FT /\*note= "base n at position 3057 is not identified in the specification"  
 FT misc\_difference 3074 /\*tag= q  
 FT /\*note= "base n at position 3074 is not identified in the specification"  
 FT WO9640950-Al.  
 FT 19-DEC-1996.  
 FT 07-JUN-1996; 96WO-US08692.  
 FT 07-JUN-1995; 95US-0476864.  
 FT (PION-) PIONEER HI-BRED INT INC.  
 FT Alhani DJ, Arnison PG, Fabijanski SF, Laurian RS;  
 FT WPI; 1997-052340/05.  
 FT P-PSDB; W08378.  
 FT Recombinant molecule for preparation of male sterile plants - used to produce stress resistant plants and for increase in seed yield  
 FT Example 2; Fig 3c(1-4); 224pp; English.

XX A fragment (T48862) is provided of microspore specific clone L16, isolated from a Brassica napus genomic library. A 2.4 kb BamHI fragment L16 contg. the coding region was added in antisense orientation to BamHI-digested PAL1007, a vector that has the CaMV 35S promoter joined to nos ter. The resulting vector, PAL1305, was used to transform Brassica napus. Plants which did not carry the kanamycin gene linked to the antisense gene died, while those that carried the antisense gene were able to flower. Thus, a constitutive promoter was used to regulate the expression of pollen-specific antisense RNA. Methods are provided for using antisense RNA to produce male sterile plants and hybrid seed.  
 XX Sequence 3293 BP; 1032 A; 701 C; 579 G; 968 T; 13 other;  
 SQ  
 Query Match 1.7%; Score 19; DB 18; Length 3293;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1097 GATAGAAAGCTGAGTCTCC 1115  
 Db 2811 gatagaagctgagtctcc 2829  
 |||||  
 RESULT 12  
 X20419/C  
 ID X20419 standard; DNA; 602 BP.  
 XX X20419;  
 XX  
 XX 04-MAY-1999 (first entry)  
 XX Human secreted protein gene 8.  
 XX  
 Human; secreted protein; fusion protein; gene therapy; protein therapy; diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia; developmental abnormality; foetal deficiency; blood; allergy; renal; ds; immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma; inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS; cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus; osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion; endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 OS Homo sapiens.  
 XX  
 XX WO9906423-Al.  
 XX 11-FEB-1999.  
 XX  
 XX 29-JUL-1998; 98WO-US15949.  
 XX 19-AUG-1997; 97US-0056730.  
 PR 30-JUL-1997; 97US-0054209.  
 PR 30-JUL-1997; 97US-0054211.  
 PR 30-JUL-1997; 97US-0054212.  
 PR 30-JUL-1997; 97US-0054213.  
 PR 30-JUL-1997; 97US-0054214.  
 PR 30-JUL-1997; 97US-0054215.  
 PR 30-JUL-1997; 97US-0054217.  
 PR 30-JUL-1997; 97US-0054218.  
 PR 30-JUL-1997; 97US-0054234.  
 PR 30-JUL-1997; 97US-0054236.  
 PR 18-AUG-1997; 97US-0055968.  
 PR 18-AUG-1997; 97US-0055969.  
 PR 18-AUG-1997; 97US-0055972.  
 PR 19-AUG-1997; 97US-0056534.  
 PR 19-AUG-1997; 97US-0056543.  
 PR 19-AUG-1997; 97US-0056554.  
 PR 19-AUG-1997; 97US-0056561.  
 PR 19-AUG-1997; 97US-0056727.  
 PR 19-AUG-1997; 97US-0056729.  
 XX

(HUMA-) HUMAN GENOME SCI INC.  
Carter KC, Endress GA, Fan P, Feng P, Kyaw H, Lafleur DW;  
Li Y, Moore PA, Rosen CA, Ruben SM, Shi Y, Wei Y;  
Zeng Z;  
WPI: 1999-153691/13.  
P-PSDB; Y00265.  
New isolated human genes and the secreted polypeptides they encode -  
useful for diagnosis and treatment of e.g. cancers, neurological  
disorders, immune diseases, inflammation or blood disorders  
Claim 1; Page 214; 312pp; English.  
This sequence represents a nucleic acid molecule which encodes a secreted  
human protein. The gene number is given in the descriptor line. The gene  
can be used to generate fusion proteins by linking to the gene to a human  
immunoglobulin Fc portion (e.g. X20403) for increasing the stability of  
the fused protein as compared to the human protein only.  
The invention relates to 83 novel genes and their fragments (nucleic acid  
sequences: X20412-X20499; amino acid sequences Y00258-Y00377) which  
are useful for preventing, treating or ameliorating medical conditions  
e.g. by protein or gene therapy. Also, pathological conditions can be  
diagnosed by determining the amount of the new polypeptides in a sample  
or by determining the presence of mutations in the new polynucleotides.  
Specific uses are described for each of the 86 polynucleotides, based on  
which tissues they are most highly expressed in (see X20412 for described  
uses).  
Sequence 602 BP; 194 A; 133 C; 152 G; 123 T; 0 other;  
Query Match 1.6%; Score 18; DB 20; Length 602;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 885 GCCTTCTTCGCACTGCTG 902  
|||||  
DB 179 GGCTTCTTCGCACTGCTG 162  
RESULT 13  
X37413/c  
ID/ X37413 standard; cDNA; 822 BP.  
X37413;  
X37413;  
06-JUL-1999 (first entry)  
Human secreted protein cDNA fragment containing gene 45.  
Human; secreted protein; prevention; treatment; protein therapy;  
gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;  
developmental abnormality; foetal deficiency; blood disorder; lymphoma;  
leukemia; immune system disorder; autoimmune disease; hepatic disease;  
renal disease; inflammation; allergy; asthma; sepsis; diabetes; AIDS;  
Alzheimer's disease; cognitive disorder; schizophrenia; osteoporosis;  
arthritis; psoriasis; digestive; endocrine; infection; ss.  
Homo sapiens.  
WO9909155-A1.  
25-FEB-1999.  
18-AUG-1998; 98WO-US17044.  
16-JUN-1998; 98US-0092956.  
15-JUL-1998; 98US-0092956.  
19-AUG-1997; 97US-0056368.  
19-AUG-1997; 97US-0056369.  
19-AUG-1997; 97US-0056535.

19-AUG-1997; 97US-0056555.  
19-AUG-1997; 97US-0056556.  
19-AUG-1997; 97US-0056628.  
19-AUG-1997; 97US-0056629.  
19-AUG-1997; 97US-0056726.  
19-AUG-1997; 97US-0056728.  
(HUMA-) HUMAN GENOME SCI INC.  
Brewer LA, Duan R, Ebner R, Endress GA, Feng P;  
Florence C, Florence KA, Komatsoulis GA, Lafleur DW;  
Moore PA, Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR;  
Young PE;  
WPI: 1999-190160/16.  
P-PSDB; Y07788.  
New isolated human genes and the secreted polypeptides they encode  
- useful for diagnosis and treatment of e.g. cancers, neurological  
disorders, immune diseases, inflammation or blood disorders  
Claim 1a; Page 216; 280pp; English.  
This invention describes novel isolated human secreted proteins and  
their encoding nucleic acid sequences. The products of the invention  
are useful for preventing, treating or ameliorating medical conditions  
e.g. by protein or gene therapy. Also pathological conditions can be  
diagnosed by determining the presence or amount of expression of  
the new polypeptides in a sample or by determining the presence or  
absence of mutations in the new polynucleotides. Specific uses are  
described for each of the 70 polynucleotides, based on which tissues they  
are most highly expressed in, and include developing products for the  
diagnosis or treatment of cancer, tumours, neurodegenerative  
disorders, developmental abnormalities and foetal deficiencies, blood  
disorders, leukemias, diseases of the immune system, autoimmune diseases,  
hepatic and renal disease, lymphomas, inflammation, allergies, asthma,  
sepsis, diabetes, Alzheimer's and cognitive disorders, schizophrenia,  
osteoporosis, arthritis, psoriasis, digestive/endocrine disorders,  
infections and AIDS. The human secreted proteins of the invention are  
represented in X37744-Y07850 and the encoding nucleic acids are  
represented in X37369-X37441.  
Sequence 822 BP; 307 A; 100 C; 136 G; 274 T; 5 other;  
Query Match 1.6%; Score 18; DB 20; Length 822;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 639 TTCTTTATCCCTTGAT 656  
|||||  
DB 116 TTCTTTATCCCTTGAT 99  
RESULT 14  
C69465/c  
ID C69465 standard; cDNA; 1092 BP.  
X37413;  
X37413;  
30-JAN-2001 (first entry)  
Human secreted protein gene 11 SEQ ID NO:21.  
Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;  
fungicide; ophthalmological; gene therapy; autoimmune disease; infection;  
hyperproliferative disorder; cardiovascular disorder; angiogenesis;  
cerebrovascular disorder; nervous system disorder; ocular disorder;  
wound healing; skin aging; food additive; preservative; ss.  
Homo sapiens.

XX PN WO200058469-A1.  
 XX XX  
 XX PD 05-OCT-2000.  
 XX PF 23-MAR-2000; 2000WO-US07579.  
 XX XX 26-MAR-1999; 99US-0126509.  
 XX PR 07-JAN-2000; 2000US-0174853.  
 XX XX  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX XX  
 XX PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX XX WPI: 2000-594642/56.  
 XX DR P-PSDB; B38213.  
 XX DR  
 XX PT Isolated nucleic acid molecule encoding a human secreted protein is  
 XX PT used in preventing, treating or ameliorating a medical condition -  
 XX XX  
 XX PS Claim 1; Page 336; 416pp; English.  
 XX XX  
 XX CC The polynucleotide sequences given in C69455 to C69502 encode the human  
 XX CC secreted proteins given in B38250 to B38320. B38251 to B38320 represent  
 XX CC human secreted polypeptide sequences and proteins homologous to them,  
 XX CC which are given in the exemplification of the present invention. Human  
 XX CC secreted proteins have activities based on the tissues and cells the  
 XX CC genes are expressed in. Example of activities include: immunosuppressive;  
 XX CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
 XX CC vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;  
 XX CC virucide; fungicide; and ophthalmological. The polynucleotides and  
 XX CC polypeptides can be used to prevent, treat or ameliorate a medical  
 XX CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 XX CC chickens or sheep. They are also used in diagnosing a pathological  
 XX CC condition or susceptibility to a pathological condition. Disorders which  
 XX CC are diagnosed or treated include autoimmune diseases, hyperproliferative  
 XX CC disorders, cardiovascular disorders, cerebrovascular disorders,  
 XX CC angiogenesis, nervous system disorders, infections caused by bacteria,  
 XX CC viruses and fungi and ocular disorders. The polypeptides can also be used  
 XX CC to aid wound healing and epithelial cell proliferation, to prevent skin  
 XX CC aging due to sunburn, to maintain organs before transplantation, for  
 XX CC supporting cell culture of primary tissues, to regenerate tissues and in  
 XX CC chemotaxis. The polypeptides can also be used as a food additive or  
 XX CC preservative to increase or decrease storage capabilities. C69446 to  
 XX CC C69454 and B38202 represent sequences used in the exemplification of the  
 XX CC present invention.

XX SQ Sequence 1092 BP; 295 A; 272 C; 298 G; 227 T; 0 other;

Query Match 1.6%; Score 18; DB 21; Length 1092;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 TCTCCCATCTGGGCAGT 284  
 |||||  
 DB 857 TCTCCCATCTGGGCAGT 840

RESULT 15  
 C98169/c  
 ID C98169 standard; cDNA; 1410 BP.  
 XX XX  
 XX AC C98169;  
 XX XX  
 XX DT 09-MAR-2001 (first entry)  
 XX XX  
 XX DE Human colon cancer antigen nucleotide sequence SEQ ID NO:179.  
 XX XX  
 XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 XX KW identification; cytostatic; cardioactive; neuroprotective; vulnary;  
 XX KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
 XX KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;

KW neural disorder; immune system disorder; muscular disorder;  
 KW reproductive disorder; gastrointestinal disorder; renal disorder;  
 KW infectious disease; cardiovascular disorder; ss.  
 XX OS Homo sapiens.  
 XX XX WO200055351-A1.  
 XX XX 21-SEP-2000.  
 XX XX 08-MAR-2000; 2000WO-US05883.  
 XX XX 12-MAR-1999; 99US-0124270.  
 XX XX (HUMA-) HUMAN GENOME SCI INC.  
 XX XX Rosen CA, Ruben SM;  
 XX XX WPI: 2000-587534/55.  
 XX XX P-PSDB; B53412.  
 XX XX  
 XX PT Colon cancer associated gene sequences, referred to as colon cancer  
 XX PT antigens, useful for the treatment, prevention, and diagnosis of colon  
 XX PT disorders such as colon cancer -  
 XX PS Claim 1; Page 604; 2104pp; English.  
 XX XX  
 XX CC C97991 to C98763 encode the human colon cancer associated proteins,  
 XX CC called human colon cancer antigens, given in B53234 to B54006. The human  
 XX CC colon cancer antigens can have cytostatic, cardioactive, muscular;  
 XX CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
 XX CC vulnary, nephrotropic, antiinfective and antibacterial activities, and  
 XX CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
 XX CC proteins and antibodies to the proteins are useful for the prevention,  
 XX CC treatment and diagnosis of colon disorders, such as colon cancer. The  
 XX CC polynucleotides may be used in diagnostics and research, such as for  
 XX CC chromosome identification, and as hybridisation probes. The proteins  
 XX CC may also be used to prevent diseases such as neural disorders, immune  
 XX CC system disorders, muscular disorders, reproductive disorders,  
 XX CC gastrointestinal disorders, wounds, renal disorders, infectious  
 XX CC diseases, and cardiovascular disorders. C98764 to C98772 and B54007  
 XX CC represent sequences used in the exemplification of the present  
 XX CC invention.

XX SQ Sequence 1410 BP; 498 A; 210 C; 247 G; 452 T; 3 other;

Query Match 1.6%; Score 18; DB 21; Length 1410;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 639 TTCTTTTATCCCTTGAT 656  
 |||||  
 DB 703 TTCTTTTATCCCTTGAT 686

RESULT 16  
 Z08776/c  
 ID Z08776 standard; DNA; 2606 BP.  
 XX XX  
 XX AC Z08776;  
 XX XX  
 XX DT 03-NOV-1999 (first entry)  
 XX XX  
 XX DE Hop lupulin regulatory DNA sequence SEQ ID NO:7.  
 XX XX  
 XX KW Hop; cone; lupulin; hop lupulin gland; dioecious; beer; bitterness;  
 XX KW secondary metabolite; genetic engineering; breeding; flavour;  
 XX KW chalcone synthetase; flavanoid biosynthesis; valerophenone synthetase;  
 XX KW bitter substance biosynthesis; food; medicine; ds.  
 XX OS Humulus lupulus.

```
PN WO942599-A1.
XX
PD 26-AUG-1999.
XX
PF 16-FEB-1999; 99WO-JP00658.
XX
XX 22-JUN-1998; 98JP-0174235.
PR 19-FEB-1998; 98JP-0037266.
XX
XX (SAPB ) SAPPORO BREWERIES LTD.
PA
PI Ito K, Okada Y;
XX
XX Claim 10; Page 37-38; 40pp; English.
XX
XX The present invention describes lupulin which is a protein that is
CC involved in secondary metabolite production in hop lupulin glands.
CC The present sequence represents a specifically claimed lupulin
CC regulatory DNA sequence given in the present invention. The lupulin
CC gene which is specifically expressed in hop lupulin glands enables the
CC breeding of hops by genetic engineering techniques focused on secondary
CC metabolites expressed in these glands. Also, vectors comprising this
CC gene can be used for the production of secondary metabolites outside of
CC plants, such as in cultured cells. Such secondary metabolites include
CC important materials such as food and drugs, and, since chalcone
CC synthetase is involved in flavanoid biosynthesis, and valerophenone
CC used for the development of food and medicines. Secondary metabolites
CC produced by the lupulin glands of hops are important, e.g. to add
CC flavour to beer. The present invention allows these to be identified and
CC expressed.
XX
SQ Sequence 2606 BP; 866 A; 568 C; 424 G; 748 T; 0 other;

Query Match 1.6%; Score 18; DB 20; Length 2606;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 TTATTTTGTGTCAGTCAA 40
Db 977 TTATTTTGTGTCAGTCAA 960
|||||
DE 208775 standard; DNA; 3439 BP.
AC 208775;
XX
XX 03-NOV-1999 (first entry)
DT
DE Hop lupulin regulatory DNA sequence SEQ ID NO:6.
XX
XX Hop; cone; lupulin; hop lupulin gland; dioecious; beer; bitterness;
KW secondary metabolite; genetic engineering; breeding; flavour;
KW chalcone synthetase; flavanoid biosynthesis; valerophenone synthetase;
KW bitter substance biosynthesis; food; medicine; ds.
XX
XX Humulus lupulus.
OS
XX WO942599-A1.
PN
XX 26-AUG-1999.
PD
XX 16-FEB-1999; 99WO-JP00658.
PF
XX 22-JUN-1998; 98JP-0174235.
PR
XX
XX (SAPB ) SAPPORO BREWERIES LTD.
PA
PI Ito K, Okada Y;
XX
XX Claim 11; Page 35-36; 40pp; English.
XX
XX The present invention describes lupulin which is a protein that is
CC involved in secondary metabolite production in hop lupulin glands.
CC The present sequence represents a specifically claimed lupulin
CC regulatory DNA sequence given in the present invention. The lupulin
CC gene which is specifically expressed in hop lupulin glands enables the
CC breeding of hops by genetic engineering techniques focused on secondary
CC metabolites expressed in these glands. Also, vectors comprising this
CC gene can be used for the production of secondary metabolites outside of
CC plants, such as in cultured cells. Such secondary metabolites include
CC important materials such as food and drugs, and, since chalcone
CC synthetase is involved in flavanoid biosynthesis, and valerophenone
CC used for the development of food and medicines. Secondary metabolites
CC produced by the lupulin glands of hops are important, e.g. to add
CC flavour to beer. The present invention allows these to be identified and
CC expressed.
XX
SQ Sequence 3439 BP; 1063 A; 790 C; 646 G; 940 T; 0 other;

Query Match 1.6%; Score 18; DB 20; Length 3439;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 TTATTTTGTGTCAGTCAA 40
Db 1667 TTATTTTGTGTCAGTCAA 1650
|||||
DE 293304 standard; DNA; 6935 BP.
AC 293304;
XX
XX 04-JUL-2000 (first entry)
DT
DE Sequence encoding human homer interacting protein I42.
XX
XX Homer; calcium; receptor; immediate early gene; IEG;
KW identification; treatment; glutamate receptor;
KW inositol triphosphate; epilepsy; glutamate toxicity;
KW memory disorder; learning disorder; stroke; brain development;
KW Alzheimer's disease; tissue degeneration; brain disorder;
KW cardiac disorder; muscular disorder; vascular disorder;
KW neurological disorder; psychiatric disorder; renal disorder;
KW uterine disorder; bronchial disorder; ageing; human; ss.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 498..4782
FT CDS /*tag= a
FT /product= Homer interacting protein I42
FT
XX WO200011204-A2.
PN
XX 02-MAR-2000.
PD
XX 18-AUG-1999; 99WO-US18973.
PF
XX
```

```

XX 18-AUG-1998; 98US-0097334.
PR 09-JUN-1999; 99US-0138426.
PR 09-JUN-1999; 99US-0138493.
PR 09-JUN-1999; 99US-0138494.
XX
PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Worley PF, Tu JC, Xiao B, Leahy D, Beneken J, Lanahan AA;
XX
XX WPI: 2000-246571/21.
DR P-PSDB; Y83008.
XX
XX Identifying compounds capable of modulating cellular response useful
PT for treating Alzheimer's disease and cardiac disorders, involves
PT incubating compound with cell expressing Homer protein and cell-surface
PT receptor .
XX
XX Disclosure; Page 139-141; 17lpp; English.
XX
XX Homer proteins are the products of neuronal immediate early genes
CC (IEG's). They selectively bind the carboxy termini of certain
CC cell-surface receptors, certain intracellular receptors and binding
CC proteins. Many forms of Homer proteins contain a 'coiled-coil'
CC structure in the carboxy terminal domain which mediated homo- and
CC heteromultimerisation between Homer proteins. Homer plays a
CC significant role in mediating receptor-activated calcium mobilisation
CC from intracellular stores. Thus, cells expressing a Homer protein
CC can be used to identify a compound capable of modulating a cellular
CC response mediated by cell surface receptor or intracellular receptor.
CC Compounds identified in this manner which modulate Homer protein
CC activity are useful for treating disorders associated with glutamate
CC receptors such as epilepsy, glutamate toxicity, memory disorders,
CC disorders of learning, stroke, schizophrenia, Alzheimer's disease,
CC tissue degeneration and disorders of brain development and also for
CC treating disorders associated with Homer protein activity which
CC includes cardiac, muscular, vascular, neurological, psychiatric,
CC renal, uterine and bronchial tissue disorders and for affecting the
CC natural aging process. These compounds are also useful for modulating
CC receptor-mediated calcium mobilization, by exposing a cell to the
CC compound to modulate calcium mobilization that normally occurs
CC when the cell is exposed to a ligand, typically an agonist or
CC antagonist of metabotropic glutamate receptors, or to activate an
CC intracellular signaling pathway, especially an inositol triphosphate
CC signaling pathway.
XX
XX Sequence 6935 BP; 2013 A; 1637 C; 1694 G; 1591 T; 0 other;
SQ

```

```

Query Match 1.6%; Score 18; DB 21; Length 6935;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 228 CTTGACATGGCCATTGC 245
|
Db 2783 cttgaacatggccattgc 2800

```

```

RESULT 19
F22286/c
ID F22286 standard; DNA; 134499 BP.
XX
AC F22286;
XX

```

```

DT 20-MAR-2001 (first entry)
XX
XX BAC containing repeats from centromeres 1-4 #9.
DE
XX Centromere; microsome; vector; ds.
XX
XX Arabidopsis thaliana.
OS
XX WO200055325-A2.
PN

```

```

XX 21-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-US07392.
XX
XX 18-MAR-1999; 99US-0125219.
PR 01-APR-1999; 99US-0127409.
PR 18-MAY-1999; 99US-0134770.
PR 13-SEP-1999; 99US-0153584.
PR 17-SEP-1999; 99US-0154603.
XX
XX (UYCH-) UNIV CHICAGO.
XX
XX Preuss D, Copenhaver G, Keith K;
XX
XX WPI: 2000-587529/55.
XX
XX Recombinant DNA construct comprising a plant centromere, useful for
PT producing stably inherited microchromosomes which can serve as vectors for
PT the construction of transgenic plant and animal cells .
XX
XX Claim 102; Page 453-484; 1449pp; English.
XX
XX The present invention relates to a recombinant DNA construct of a plant
CC (Arabidopsis thaliana) centromere. The constructs are useful for
CC producing stably inherited microchromosomes which can serve as vectors for
CC the construction of transgenic plant and animal cells expressing
CC selected proteins such as hormones, enzymes, interleukins, clotting
CC factors, cytokines, antibodies, and growth factors.
XX
XX Sequence 134499 BP; 41565 A; 25130 C; 25225 G; 42577 T; 2 other;
SQ

```

```

Query Match 1.6%; Score 18; DB 21; Length 134499;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 731 TCCGTGTAATCATAGCTG 748
|
Db 73979 TCCGTGTAATCATAGCTG 73962

```

```

RESULT 20
V21209/c
ID V21209 standard; DNA; 1664976 BP.
XX
AC V21209;
XX

```

```

DT 10-NOV-1998 (first entry)
XX
XX Methanococcus jannaschii circular chromosome.
DE
XX
XX Methanococcus jannaschii; methanogenic archaeon; circular chromosome;
KW genome; autotrophic; extrachromosomal element; identification; ds.
XX
XX Methanococcus jannaschii.
OS
XX
XX WO9807830-A2.
XX

```

```

PD 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-US14900.
PF
XX
XX 22-AUG-1996; 96US-0024428.
PR
XX
XX (GENO-) INST GENOMIC RES.
PA
XX (UNII ) UNIV ILLINOIS FOUND.
PA
XX (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
XX Bult CJ, Smith HO, Venter JC, White OR, Woese CR;
XX
XX WPI: 1998-169145/15.
XX

```

PT Complete genome sequence of methano-genic archaeon, Methanococcus  
PT jannaschii - useful in identification of M. jannaschii genome  
PT fragment  
XX  
XX  
XX Claim 13; Page 152-585; 614pp; English.  
XX The present sequence represents the complete 1.66-megabase pair genome  
CC sequence of the Methanococcus jannaschii circular chromosome. The  
CC present invention describes M. jannaschii open reading frames from the  
CC genome sequence. The invention also describes a computer based system  
CC for identifying fragments of the M. jannaschii genome that are  
CC homologous to target nucleotide sequences, comprising: (a) data storage  
CC means comprising the nucleotide sequence of the 1664976, 58407 or 16550  
CC bp sequence (see V21209, V21210 and V21211), or a nucleotide sequence at  
CC least 99.9% identical to it; (b) search means for comparing a target  
CC sequence to the nucleotide sequence of the data storage means to  
CC identify a homologous sequence, and (c) retrieval means for obtaining  
CC the homologous sequence. The method, which is based on whole genome  
CC random sequencing of an autotrophic archaeon M. jannaschii, the genome  
CC of which consists of 3 physically distinct elements, a large circular  
CC chromosome (the 1664976 bp sequence given in V21209), a large circular  
CC extra-chromosomal element (the 58407 bp sequence given in V21210), and a  
CC small circular extra-chromosomal element (the 16550 bp sequence given in  
CC V21211), can be used in the identification of M. jannaschii genome  
CC fragment.  
XX  
XX Sequence 1664976 BP; 568133 A; 264649 C; 258701 G; 573392 T; 101 other;  
SQ

Query Match 1.6%; Score 18; DB 19; Length 1664976;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 CTTGATGCTTCATGATAT 667  
|||||  
Db 779525 CTTGATGCTTCATGATAT 779508

RESULT 21  
Z14869/c  
ID Z14869 standard; cDNA; 300 BP.  
XX  
XX Z14869;  
XX  
XX 12-OCT-1999 (first entry)  
XX Human gene expression product cDNA sequence SEQ ID NO:2338.  
XX  
XX Human; gene; gene expression product; diagnosis; therapy; probe;  
KW detection; mapping; tissue typing; profiling; forensic; cancer;  
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO9938972-A2.  
XX  
XX 05-AUG-1999.  
XX  
XX 28-JAN-1999; 99WO-US01619.  
XX  
XX 03-APR-1998; 98US-0080666.  
XX 24-JAN-1998; 98US-0072910.  
XX 24-FEB-1998; 98US-0075954.  
XX 31-MAR-1998; 98US-0080114.  
XX 03-APR-1998; 98US-0080515.  
XX  
XX (CHIR ) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
XX  
XX Crikvenjakov R, Drmanac M, Drmanac R, Drmanac S;  
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;  
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;

PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
XX WPI; 1999-494092/41.  
XX  
XX Novel human genes and their expression products which are  
PT differentially expressed in different cell types  
PT  
XX  
XX Claim 1; Page 1186; 2479pp; English.  
XX The present invention describes a library of human polynucleotides  
CC comprising the sequences given in 212532 to 217779. Also described is a  
CC method of detecting differentially expressed genes correlated with the  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample from a cell  
CC suspected of being cancerous, where the gene product is encoded by one  
CC of the 5248 polynucleotide sequences given in 212532 to 217779. The  
CC polynucleotides can be used as a source of primers and probes, which can  
CC be used for a variety of purpose, e.g. detection of expression levels,  
CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein); and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.  
XX  
XX Sequence 300 BP; 90 A; 58 C; 51 G; 101 T; 0 other;  
SQ

Query Match 1.5%; Score 17; DB 20; Length 300;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 754 CTTGTGTTCTCGCTTG 770  
|||||  
Db 251 CTTGTGTTCTCGCTTG 235

RESULT 22  
C03809/c  
ID C03809 standard; cDNA; 380 BP.  
XX  
XX C03809;  
XX  
XX 06-OCT-2000 (first entry)  
XX Human secreted protein 5' EST, SEQ ID NO: 3807.  
XX  
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
XX Homo sapiens.  
XX  
XX EP1033401-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 21-FEB-2000; 2000EP-0200610.  
XX  
XX 26-FEB-1999; 99US-0122487.  
XX  
XX (GEST ) GENSET.  
XX  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
PI WPI: 2000-500381/45.  
XX P-PSDB; G03803.  
XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures  
 XX  
 PS Claim 1; SEQ ID 3807; 71pp + CD-ROM; English.  
 XX  
 CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. An ORF has been identified within the  
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
 CC derived from 30 different tissues. EST sequences usually correspond  
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
 CC well suited for isolating cDNA sequences derived from the 5' ends of  
 CC mRNAs and even in those cases where longer cDNA sequences have been  
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
 CC gene therapy and chromosome mapping procedures. They are used to obtain  
 CC upstream regulatory sequences and to design expression and secretion  
 CC vectors.  
 XX  
 SQ Sequence 380 BP; 77 A; 88 C; 118 G; 97 T; 0 other;

Query Match 1.5%; Score 17; DB 21; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 892 CTGCACCTGCTCCCTGAA 908  
 Db 108 CTGCACCTGCTCCCTGAA 92  
 |||||

## RESULT 23

A80659  
 ID A80659 standard; cDNA; 478 BP.

XX A80659;

XX 21-NOV-2000 (first entry)

XX Human secreted protein gene #7 related gene HNHHA34R SEQ ID #111.

XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;  
 KW antirheumatic, dermatological; antiproliferative; antiarteriosclerotic;  
 KW anticancer; vulnery; antiviral; antibacterial; antifungal;  
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;  
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;  
 KW Crohn's disease; nephritis; hyperproliferative disorder;  
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;  
 KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23; ss.

XX Homo sapiens.

XX WO200029435-A1.

XX 25-MAY-2000.

XX 27-OCT-1999; 99WO-US25031.

XX 28-OCT-1998; 98US-0105971.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;  
 PI Greene JM;

XX WPI; 2000-387742/33.

XX Isolated nucleic acid molecules encoding human secreted proteins are  
 PT used for the prevention, amelioration and treatment of autoimmune,  
 PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,  
 PT wounds, and infectious diseases.

XX Disclosure; Page 754; 803pp; English.

XX The present invention relates to 12 secreted human proteins and the  
 CC nucleotide sequences encoding them. The polynucleotide sequences given in  
 CC A80606-A80623 encode the 12 secreted protein sequences given in  
 CC B25576-B25593. The human secreted proteins have various activities  
 CC dependent on the tissues in which they are expressed. Examples of the  
 CC activities of the proteins include: immunosuppressant; anti-inflammatory;  
 CC antiarthritic; antirheumatic, dermatological; antiproliferative;  
 CC antiarteriosclerotic; anticancer; vulnery; antiviral; antibacterial;  
 CC and antifungal activity. The proteins, polypeptides, agonists and  
 CC antagonists may be used to treat prevent and/or diagnose various disease,  
 CC disorders and conditions examples of which include: immune disorders  
 CC e.g. Addison's disease, rheumatoid arthritis, dermatitis, and multiple  
 CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,  
 CC Crohn's disease and nephritis; hyperproliferative disorders such as  
 CC paraneoplasias and purpura; cardiovascular disorders e.g. coronary  
 CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The  
 CC proteins and polynucleotide sequences may also be used in wound healing  
 CC and the treatment of infectious diseases.  
 CC The human secreted protein gene #7 and protein sequences are represented  
 CC in sequences A80612 and B25582. Secreted protein gene #7 is located at  
 CC position chromosome 15 q22.3-23. Sequences A80652-A80661 represent genes  
 CC which are related to the secreted protein gene#7.

XX Sequence 478 BP; 105 A; 133 C; 124 G; 108 T; 8 other;

Query Match 1.5%; Score 17; DB 21; Length 478;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 883 CTGGCTTCTCCTGACATG 899

Db 191 ctggcttctcgcactg 207  
 |||||

## RESULT 24

XX X14062  
 ID X14062 standard; DNA; 686 BP.

XX X14062;

XX 31-MAR-1999 (first entry)

XX H. pylori GHPO 454 gene.

XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease; ss.

XX Helicobacter pylori.

XX Key Location/Qualifiers  
 FT CDS 82..636  
 FT /\*tag= a

XX WO9843478-A1.

XX 08-OCT-1998.

XX 01-APR-1998; 98WO-US06371.

XX 29-JUL-1997; 97US-0902615.

XX 01-APR-1997; 97US-0833457.

XX 24-JUN-1997; 97US-0881227.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (INNR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
 DR WPI; 1998-542293/46.

DR P-PSDB; W98343.  
 XX New isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 PT infections and gastrointestinal diseases  
 XX  
 XX  
 PS Claim 1; Page 491-492; 2054pp; English.  
 XX  
 CC This sequence represents a polynucleotide of the invention. It was  
 CC isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.  
 CC The polypeptides can be used for preventing or treating Helicobacter  
 CC infections, and gastroduodenal diseases associated with these  
 CC infections, including acute, chronic, and atrophic gastritis, and peptic  
 CC ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used  
 CC for the production of antibodies. The products can also be used for  
 CC detection and diagnosis.  
 XX  
 SQ Sequence 686 BP; 220 A; 110 C; 148 G; 208 T; 0 other;  
 Query Match 1.5%; Score 17; DB 19; Length 686;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 914 TGCTCTACGCTTTTATT 930  
 DB 670 tgctctacgctttttatt 686  
 RESULT 25  
 X14375/c  
 ID X14375 standard; DNA; 756 BP.  
 XX AC X14375;  
 XX  
 XX 31-MAR-1999 (first entry)  
 XX  
 DE H. pylori GHPO 453 gene.  
 XX  
 XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease; ss.  
 XX  
 XX Helicobacter pylori.  
 OS  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 36..672  
 FT /\*tag= a  
 XX  
 XX W09843478-A1.  
 PN  
 XX  
 PD 08-OCT-1998.  
 XX  
 XX 01-APR-1998; 98WO-US06371.  
 XX  
 XX 29-JUL-1997; 97US-0902615.  
 PR 01-APR-1997; 97US-0833457.  
 PR 24-JUN-1997; 97US-0881227.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 XX  
 XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
 PI  
 XX WPI; 1998-542293/46.  
 DR P-PSDB; W98656.  
 XX  
 XX New isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 PT infections and gastrointestinal diseases  
 XX  
 XX Claim 1; Page 1344-1345; 2054pp; English.  
 PS  
 XX This sequence represents a polynucleotide of the invention. It was

CC isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.  
 CC The polypeptides can be used for preventing or treating Helicobacter  
 CC infections, and gastroduodenal diseases associated with these  
 CC infections, including acute, chronic, and atrophic gastritis, and peptic  
 CC ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used  
 CC for the production of antibodies. The products can also be used for  
 CC detection and diagnosis.  
 XX  
 SQ Sequence 756 BP; 244 A; 116 C; 178 G; 218 T; 0 other;  
 Query Match 1.5%; Score 17; DB 19; Length 756;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 914 TGCTCTACGCTTTTATT 930  
 DB 661 TGCTCTACGCTTTTATT 645  
 RESULT 26  
 Z15412/c  
 ID Z15412 standard; cDNA; 768 BP.  
 XX AC Z15412;  
 XX  
 XX 12-OCT-1999 (first entry)  
 DT  
 XX  
 DE Human gene expression product cDNA sequence SEQ ID NO:2881.  
 XX  
 XX Human; gene; gene expression product; diagnosis; therapy; probe;  
 KW detection; mapping; tissue typing; profiling; forensic; cancer;  
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX W09938972-A2.  
 PN  
 XX  
 XX 05-AUG-1999.  
 PD  
 XX  
 XX 28-JAN-1999; 99WO-US01619.  
 PF  
 XX  
 XX 03-APR-1998; 98US-0080566.  
 PR 28-JAN-1998; 98US-0072910.  
 PR 24-FEB-1998; 98US-0075954.  
 PR 31-MAR-1998; 98US-0080114.  
 PR 03-APR-1998; 98US-0080515.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX  
 PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;  
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;  
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
 XX  
 XX WPI; 1999-494092/41.  
 DR  
 XX  
 XX Novel human genes and their expression products which are  
 PT differentially expressed in different cell types  
 PT  
 XX  
 XX Claim 1; Page 1392; 2479pp; English.  
 PS  
 XX  
 CC The present invention describes a library of human polynucleotides  
 CC comprising the sequences given in Z12532 to Z17779. Also described is a  
 CC method of detecting differentially expressed genes correlated with the  
 CC cancerous state of a mammalian cell, comprising detecting at least one  
 CC differentially expressed gene product in a test sample from a cell  
 CC suspected of being cancerous, where the gene product is encoded by one  
 CC of the 5248 polynucleotide sequences given in Z12532 to Z17779. The  
 CC polynucleotides can be used as a source of primers and probes, which can  
 CC be used for a variety of purpose, e.g. detection of expression levels,

CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein); and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.  
XX

SQ Sequence 768 BP; 206 A; 155 C; 149 G; 156 T; 102 other;

Query Match 1.5%; Score 17; DB 20; Length 768;

Best Local Similarity 100.0%; Pred. No. 75;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GATGTCATCATATTTT 670

|||||

Db 294 GATGTCATCATATTTT 278

RESULT 27

C36461/C

ID C36461 standard; DNA; 822 BP.

XX AC C36461;

XX DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 13892.

KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 16-APR-1999; 99US-0128714.

XX PR 19-APR-1999; 99US-0129845.

XX PR 21-APR-1999; 99US-0130077.

XX PR 23-APR-1999; 99US-0130449.

XX PR 28-APR-1999; 99US-0130510.

XX PR 30-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132487.

XX PR 11-MAY-1999; 99US-0132863.

XX PR 14-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161408.  
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 1.5%; Score 17; DB 21; Length 822;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 767 CTGTCAGATTCTTCAT 783  
Db 513 CTGTCAGATTCTTCAT 497  
|||||  
  
RESULT 28  
C35528/c  
ID C35528 standard; DNA; 861 BP.  
XX  
AC C35528;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 10519.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137538.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139432.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 21-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 06-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.

```
PR 29-OCT-1999; 99US-0162142.
Query Match 1.5%; Score 17; DB 21; Length 861;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 767 CTTGTCAGATTCCTCAT 783
Db 551 CTTGTCAGATTCCTCAT 535
RESULT 29
C53013
ID C53013 standard; DNA; 982 BP.
XX C53013;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 73111.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
XX -06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0138817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
```

PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161923.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 17; DB 21; Length 982;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 TTGATCTGTGTTCTTGG 152

Db 942 ttgatctgtgtcttgg 958  
|||||  
RESULT 30  
Q47197/c  
ID Q47197 standard; DNA; 1032 BP.  
XX  
AC Q47197;  
XX  
DT 26-JAN-1994 (first entry)  
XX  
DE DNA polymerase III holoenzyme delta subunit hola gene.  
XX  
KW Enzyme; chromosomal replicase; ss.  
XX  
OS Escherichia coli.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1032  
FT /\*tag= a  
XX  
PN WO9315115-A.  
XX  
PD 05-AUG-1993.  
XX  
PF 22-JAN-1993; 93WO-US00627.  
XX  
PR 24-JAN-1992; 92US-0826926.  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
PI O'Donnell ME;  
XX  
DR WPI; 1993-258618/32.  
DR P-PSDB; R40125.  
XX  
PT Nucleotide sequences for DNA polymerase III holo enzyme sub-units  
PT - enzyme with 5 or 6 sub-units having same activity as 10 subunit  
PT enzyme  
XX  
PS Example; Page 15-16; 115pp; English.  
XX  
CC The sequence is that of the DNA polymerase III holoenzyme delta  
CC subunit hola gene.  
XX  
SQ Sequence 1032 BP; 241 A; 262 C; 278 G; 251 T; 0 other;

Query Match 1.5%; Score 17; DB 14; Length 1032;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 880 GTCCTGCTTTCCTGCA 896  
|||||  
Db 108 GTCCTGCTTTCCTGCA 92

RESULT 31  
T98154/c  
ID T98154 standard; DNA; 1032 BP.  
XX  
AC T98154;  
XX  
DT 16-MAR-1998 (first entry)  
XX  
DE E. coli DNA polymerase III delta subunit coding sequence.  
XX

KW Delta subunit; E. coli; DNA polymerase III holoenzyme; probe; enzyme;  
KW hybridisation; PCR; ss.  
XX  
OS Escherichia coli.  
XX

```

FH Key      Location/Qualifiers
FT CDS      1..1032
FT          /*tag= a
FT          /gene= "hola"
FT          /product= "DNA polymerase III delta subunit"
FT          /transl_except= (pos: 538..540, aa: Asn)
FT          /transl_except= (pos: 715..717, aa: Gly)
XX
XX US5668004-A.
XX
XX 16-SEP-1997.
XX
XX 22-JUL-1994; 94US-0279058.
XX
XX 22-JUL-1994; 94US-0279058.
XX 22-JAN-1993; 93US-0826926.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX O'donnell ME;
XX
XX WPI; 1997-470038/43.
XX P-PSDB; W36070.
XX
XX Isolated theta, chi, psi, delta and delta' protein sub:units of E.
XX coli polymerase III holo:enzyme - used to make man-made enzyme of 5
XX PT or 6 sub:units, useful for long chain PCR
XX
XX Claim 31; Column 9-12; 65pp; English.
XX
XX This is the nucleotide sequence encoding the delta subunit of the E. coli
XX DNA polymerase III holoenzyme. The sequence was isolated from a lambda
XX phage library using probes based on the sequences of the tryptic peptides
XX W36072-W36075.
XX The E. coli polymerase III subunits (theta, chi, psi, delta and delta',
XX (T98151-T98155 respectively)) are used to make man-made enzymes
XX comprising 5 or 6 subunits and potentially for use in long chain PCR.
XX
XX Sequence 1032 BP; 241 A; 262 C; 278 G; 251 T; 0 other;

Query Match      1.5%; Score 17; DB 18; Length 1032;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 880 GTCCTGGCTTTCCTGCA 896
Db 108 GTCCTGGCTTTCCTGCA 92

RESULT 32
T85163
ID T85163 standard; cDNA; 1059 BP.
XX
XX T85163;
XX
XX 14-DEC-1997 (first entry)
XX
XX Macaque chemokine receptor 88C.
XX
XX Chemokine receptor 88C; atherosclerosis; rheumatoid arthritis;
XX tumour; asthma; viral infection; AIDS; inflammation;
XX autoimmune disease; therapy; diagnosis; leukocyte trafficking;
XX G protein coupled receptor; ss.
XX
XX Macaca sp.
XX
XX W09722698-A2.
XX
XX 26-JUN-1997.
XX
XX 20-DEC-1996; 96WO-US20759.
XX

```

```

PR 07-JUN-1996; 96US-0661393.
PR 20-DEC-1995; 95US-0575967.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW, Raport CJ, Schweickart VL;
XX
XX WPI; 1997-341689/31.
XX P-PSDB; W27125.
XX
XX New nucleic acid encoding chemokine receptors 88-2B and 88C - used
XX to modulate leukocyte trafficking, e.g. for treatment of
XX inflammation, tumours, viral infections, autoimmune diseases, etc.
XX
XX Claim 47; Page 56-57; 65pp; English.
XX
XX This sequence comprises a coding sequence for macaque chemokine
XX receptor 88C (W27125). It was isolated from macaque genomic DNA
XX using primers (see T85176-77) based on human 88C cDNA (see T85161).
XX The macaque 88C nucleotide sequence is 98% identical to that of
XX human 88C. 88C cDNAs can be used to produce recombinant
XX polypeptides in transformed host cells for use in the treatment of
XX e.g. atherosclerosis, rheumatoid arthritis, tumours, asthma, viral
XX infection, AIDS and inflammatory conditions. Nucleic acid
XX fragments can be used to isolate genomic sequences, to detect
XX alleles of the gene (for diagnosis or in gene therapy), to alter
XX receptor genetics to facilitate identification of modulators and to
XX produce knockout animals, and (antisense forms) to alter/study the
XX genetics and expression of the receptor.
XX
XX Sequence 1059 BP; 249 A; 274 C; 227 G; 309 T; 0 other;

Query Match      1.5%; Score 17; DB 18; Length 1059;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951
Db 905 agaagtcagaaactac 921

RESULT 33
V23992
ID V23992 standard; DNA; 1059 BP.
XX
XX V23992;
XX
XX 06-AUG-1998 (first entry)
XX
XX Human CC-CKR5 coding sequence.
XX
XX CC-CKR5; chemokine receptor; mouse; human; transgenic mouse;
XX HIV infection; T-cell mediated inflammation; ss.
XX
XX Homo sapiens.
XX
XX EP834564-A2.
XX
XX 08-APR-1998.
XX
XX 03-OCT-1997; 97EP-0307823.
XX
XX 03-OCT-1996; 96US-0724984.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Bergsma DJ, Brawnner ME, Shabon U;
XX
XX WPI; 1998-195463/18.
XX
XX New isolated mouse chemokine receptor, CC-CKR5 - used to develop
XX products for the study, diagnosis and treatment of HIV infection or

```

PT T-cell mediated inflammation  
 PS Disclosure; Page 22-23; 27pp; English.

XX This sequence encodes the human CC-CR5 protein. The mouse version of  
 CC this CC-CR5 coding sequence is the DNA of the invention. CC-CR5 is a  
 CC chemokine receptor. Cells transformed with the DNA can be cultivated and  
 CC the expression product harvested. The DNA can be knocked out and replaced  
 CC with the human CC-CR5 gene, creating transgenic mice which can be used  
 CC in the study of HIV infection or T-cell mediated inflammation. Transgenic  
 CC mice could also be used to screen for human CC-CR5 agonists or  
 CC antagonists.

XX Sequence 1059 BP; 248 A; 271 C; 228 G; 312 T; 0 other;

Query Match 1.5%; Score 17; DB 19; Length 1059;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 935 AGAAGTTCAGAACTAC 951  
 Db 905 agaagttcagaactac 921  
 |||||

# RESULT 34

V84125  
 ID V84125 standard; cDNA; 1071 BP.

XX V84125;

DT 15-MAR-1999 (first entry)

XX HIV-1 co-receptor CCR5 variant CCR5m303 cDNA.

KW HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;  
 KW gene therapy; human; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 7..309  
 FT /\*tag= a

XX WO9854317-A1.

XX 03-DEC-1998.

XX 29-MAY-1998; 98WO-EP03437.

XX 30-MAY-1997; 97US-0048057.

XX (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.

XX Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;

XX WPI; 1999-059835/05.

XX P-PSDB; W88231.

XX New CCR5 variant protein of the HIV-1 co-receptor - useful in  
 XX developing resistance of CCR5-expressing cells to HIV-1 infection

XX Claim 5; Page 36-37; 55pp; English.

XX This nucleotide sequence includes a coding region for a CCR5  
 CC variant protein (see W88231), designated CCR5m303, comprising the  
 CC first two transmembrane domains of wild-type CCR5 (see W88232), but  
 CC lacking transmembrane domains 3-7. CCR5 serves as a co-receptor  
 CC for infection by macrophage-tropic (M-tropic) strains of HIV-1.  
 CC The presence of the CCR5m303 variant with the wild type CCR5  
 CC allele shows a positive correlation with resistance to infection  
 CC with M-tropic HIV-1 strains, and may indicate slower progression of  
 CC the disease. The detection of CCR5 variants may be used to

CC identify individuals at lower risk of infection relative to the  
 CC general population who, if infected, may exhibit slower progression  
 CC to AIDS. Probes and primers (see V84127-36) are provided for use  
 CC in diagnostic methods for detecting the presence of such variants.  
 CC A method is provided for inhibiting HIV-1 infection of a cell  
 CC expressing the CCR5 receptor. This involves introducing a nucleic  
 CC acid encoding a CCR5 variant into the cell, thereby reducing the  
 CC number of functional CCR5 molecules present on the cell surface.

XX Sequence 1071 BP; 253 A; 273 C; 230 G; 315 T; 0 other;

Query Match 1.5%; Score 17; DB 20; Length 1071;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 935 AGAAGTTCAGAACTAC 951  
 Db 911 agaagttcagaactac 927  
 |||||

# RESULT 35

T76920

ID T76920 standard; cDNA; 1225 BP.

XX T76920;

XX 08-JUN-1998 (first entry)

XX DNA encoding human CC chemokine receptor 5 (CCR5).

XX CC chemokine receptor 5; CCR5; G-protein coupled receptor;  
 KW human immunodeficiency virus; HIV; CD4; AIDS; therapy;  
 KW transgenic animal; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 27..1085  
 FT /\*tag= a

XX WO9745543-A2.

XX 04-DEC-1997.

XX 28-MAY-1997; 97WO-US09586.

XX 28-MAY-1996; 96US-0018508.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Alkhatib G, Berger EA, Broder CC, Combadiere C;  
 XX Feng Y, Kennedy PE, Murphy PM;

XX WPI; 1998-032650/03.

XX P-PSDB; W23835.

XX CC chemokine receptor 5 polypeptide - used to inhibit membrane  
 XX fusion between HIV and a target cell

XX Claim 15; Fig 1C; 70pp; English.

XX This DNA sequence codes for a novel human macrophage-selective CC  
 CC chemokine receptor (see W23835) designated CCR5. It was isolated  
 CC from a lambda gt11 cDNA library. Another isolated DNA molecule (see  
 CC T76919) codes for an Alai27Leu variant (see W23834) of CCR5. The  
 CC susceptibility of human macrophages to HIV infection depends on  
 CC cell surface expression of the human CD4 molecule and CCR5. CCR5  
 CC is a member of the 7-transmembrane superfamily of G-protein  
 CC coupled cell surface molecules. It plays an essential role in the  
 CC membrane fusion step of infection by some HIV isolates. The  
 CC establishment of stable, non-human cell lines and transgenic  
 CC mammals having cells that coexpress human CD4 and CCR5 provides

CC valuable tools for research of HIV infection.. Antibodies that bind  
 CC to CCR5, CCR5 variants, and CCR5-binding agents capable of blocking  
 CC membrane fusion between HIV and target cells represent potential  
 CC anti-HIV therapeutics for macrophage tropic strains of HIV.  
 CC Antisense oligonucleotides can also be used to inhibit expression  
 CC of CCR5 in a cell.

XX Sequence 1225 BP; 289 A; 302 C; 280 G; 354 T; 0 other;  
 SQ

Query Match 1.5%; Score 17; DB 19; Length 1225;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951  
 |||||  
 Db 931 agaagttcagaaactac 947

## RESULT 36

T76919  
 ID T76919 standard; cDNA; 1255 BP.

XX AC  
 XX T76919;

DT 08-JUN-1998 (first entry)

DE DNA encoding human CC chemokine receptor 5 (CCR5) Al27L variant.

XX CC chemokine receptor 5; CCR5; G-protein coupled receptor;  
 KW human immunodeficiency virus; HIV; CD4; AIDS; therapy;  
 KW transgenic animal; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FH 1..1116  
 FT CDS /\*tag= a  
 FT

XX W09745543-A2.

XX 04-DEC-1997.

XX 28-MAY-1997; 97WO-US09586.

XX 28-MAY-1996; 96US-0018508.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Alkhatib G, Berger EA, Broder CC, Combadiere C;  
 PI Feng Y, Kennedy PE, Murphy PM;

XX WPI; 1998-032650/03.  
 DR P-PSDB; W23834.

XX CC chemokine receptor 5 polypeptide - used to inhibit membrane  
 PT fusion between HIV and a target cell

XX Claim 15; Fig 1B; 70pp; English.

XX This DNA sequence codes for an Alal27Leu variant (see W23834)  
 CC of a novel human macrophage-selective CC chemokine receptor (see  
 CC also W23835) that has been designated CCR5. It was isolated from a  
 CC lambda gt11 cDNA library. The conservative amino acid substitution  
 CC should not affect the biological activity of CCR5. The  
 CC susceptibility of human macrophages to HIV infection depends on  
 CC cell surface expression of the human CD4 molecule and CCR5. CCR5  
 CC is a member of the 7-transmembrane superfamily of G-protein  
 CC coupled cell surface molecules. It plays an essential role in the  
 CC membrane fusion step of infection by some HIV isolates. The  
 CC establishment of stable, non-human cell lines and transgenic  
 CC mammals having cells that coexpress human CD4 and CCR5 provides  
 CC valuable tools for research of HIV infection. Antibodies that bind

CC to CCR5, CCR5 variants, and CCR5-binding agents capable of blocking  
 CC membrane fusion between HIV and target cells represent potential  
 CC anti-HIV therapeutics for macrophage tropic strains of HIV.  
 CC Antisense oligonucleotides can also be used to inhibit expression  
 CC of CCR5 in a cell.

XX Sequence 1255 BP; 295 A; 311 C; 289 G; 360 T; 0 other;  
 SQ

Query Match 1.5%; Score 17; DB 19; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951  
 |||||  
 Db 962 agaagttcagaaactac 978

## RESULT 37

F32719/c

ID F32719 standard; cDNA; 1272 BP.

XX AC  
 XX F32719;

DT 23-MAR-2001 (first entry)

DE Human secreted protein gene 21 SEQ ID NO:31.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 KW cerebroprotective; neutropic; neuroprotective; antibacterial; virucide;  
 KW fungicide; ophthalmological; autoimmune disease; rheumatoid arthritis;  
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 KW cardiac arrest; cerebrovascular disorder; cerebral ischaemia; infection;  
 KW angiogenesis; nervous system disorder; Alzheimer's disease; skin aging;  
 KW ocular disorder; corneal infection; wound healing; food additive;  
 KW preservative; ss.

XX Homo sapiens.

XX W0200077255-A1.

XX 21-DEC-2000.

XX 01-JUN-2000; 2000WO-US14926.

XX 11-JUN-1999; 99US-0138628.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis GA;

XX WPI; 2001-025337/03.

XX P-PSDB; B64442.

XX Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -

XX Claim 1; Page 490; 593pp; English.

XX The polynucleotide sequences given in F32699 to F32747 encode the human  
 CC secreted proteins given in B64422 to B64470. B64471 to B64548 represent  
 CC human secreted polypeptide sequences and proteins homologous to them,  
 CC which are given in the exemplification of the present invention. Human  
 CC secreted proteins have activities based on the tissues and cells the  
 CC genes are expressed in. Examples of activities include: antiarthritic;  
 CC immunosuppressive; antirheumatic; antiproliferative; cytostatic;  
 CC cardiant; vasotropic; cerebroprotective; neutropic; neuroprotective;  
 CC antibacterial; virucide; fungicide; and ophthalmological. The  
 CC polynucleotides and polypeptides can be used to prevent, treat or  
 CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,  
 CC horses, cats, dogs, chickens or sheep. They are also used in diagnosing  
 CC a pathological condition or susceptibility to a pathological condition.

Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. F32690 to F32698 and B64421 represent sequences used in the exemplification of the present invention.

Sequence 1272 BP; 318 A; 358 C; 257 G; 334 T; 5 other;

Query Match 1.5%; Score 17; DB 22; Length 1272;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 979 TGTGTGAGAGGAGTA 995

|||||

Db 488 TGTGTGAGAGGAGTA 472

RESULT 38

V84159

ID V84159 standard; cDNA; 1344 BP.

XX AC V84159;

XX DT 15-MAR-1999 (first entry)

XX DE HIV-1 co-receptor CCR5-delta32 cDNA.

XX KW HIV-1; CCR5; CCR5-delta32; co-receptor; infection; diagnosis; AIDS;

XX KW gene therapy; human; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 240..887

FT /\*tag= a

XX WO9854317-A1.

XX PD 03-DEC-1998.

XX PF 29-MAY-1998; 98WO-EP03437.

XX PR 30-MAY-1997; 97US-0048057.

XX PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.

XX PI Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;

XX DR WPI; 1999-059835/05.

XX DR P-PSDB; W88238.

XX PT New CCR5 variant protein of the HIV-1 co-receptor - useful in

XX PT developing resistance of CCR5-expressing cells to HIV-1 infection

XX PS Disclosure; Page 38-39; 55pp; English.

XX CC This nucleotide sequence includes a coding region for a CCR5

CC variant (see W88238), designated CCR5-delta32, that contains a

CC 32 bp deletion when compared to the wild-type allele (see V84126)

CC which results in a protein having a truncation of C-terminal

CC residues 188-352 (including the 5th, 6th and 7th transmembrane

CC domains). Individuals homozygous for this mutation are resistant

CC to HIV-1 infection. Heterozygous individuals, however, are

CC susceptible. The invention additionally relates to the

identification of another CCR5 variant, CCR5m303 (see W88231), lacking transmembrane domains 3-7 of CCR5. The detection of CCR5 variants may be used to identify individuals at lower risk of infection relative to the general population who, if infected, may exhibit slower progression to AIDS. Probes and primers (see W84127-36) are provided for use in diagnostic methods for detecting the presence of such variants. A method is provided for inhibiting HIV-1 infection of a cell expressing the CCR5 receptor. This involves introducing a nucleic acid encoding a CCR5 variant into the cell, thereby reducing the number of functional CCR5 molecules present on the cell surface.

Sequence 1344 BP; 336 A; 332 C; 291 G; 385 T; 0 other;

Query Match 1.5%; Score 17; DB 20; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951

|||||

Db 1112 agaagttcagaactac 1128

RESULT 39

V84126

ID V84126 standard; cDNA; 1376 BP.

XX AC V84126;

XX DT 15-MAR-1999 (first entry)

XX DE HIV-1 co-receptor CCR5 cDNA.

XX KW HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;

XX KW gene therapy; human; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 240..1298

FT /\*tag= a

FT misc\_difference 540..542

FT /\*tag= b

FT /\*note= "is TGT (Cys) in wild-type allele, TGA

FT (stop) in CCR5m303 allele"

XX WO9854317-A1.

XX PD 03-DEC-1998.

XX PF 29-MAY-1998; 98WO-EP03437.

XX PR 30-MAY-1997; 97US-0048057.

XX PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.

XX PI Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;

XX DR WPI; 1999-059835/05.

XX DR P-PSDB; W88232.

XX PT New CCR5 variant protein of the HIV-1 co-receptor - useful in

XX PT developing resistance of CCR5-expressing cells to HIV-1 infection

XX PS Claim 5; Page 32-34; 55pp; English.

XX CC This nucleotide sequence includes a coding region for CCR5 (see

CC W88231), which serves as a co-receptor for infection by

CC macrophage-tropic (M-tropic) strains of HIV-1. The invention

CC relates to the identification of a CCR5 variant (see W88231),

CC designated CCR5m303, comprising the first two transmembrane domains

CC of wild-type CCR5, but lacking transmembrane domains 3-7. The

CC presence of the CCR5n303 variant with the wild type CCR5 allele  
 CC shows a positive correlation with resistance to infection with  
 CC M-tropic HIV-1 strains, and may indicate slower progression of the  
 CC disease. The detection of CCR5 variants may be used to identify  
 CC individuals at lower risk of infection relative to the general  
 CC population who, if infected, may exhibit slower progression to  
 CC AIDS. Probes and primers (see V84127-36) are provided for use in  
 CC diagnostic methods for detecting the presence of such variants. A  
 CC method is provided for inhibiting HIV-1 infection of a cell  
 CC expressing the CCR5 receptor. This involves introducing a nucleic  
 CC acid encoding a CCR5 variant into the cell, thereby reducing the  
 CC number of functional CCR5 molecules present on the cell surface.  
 XX  
 SQ Sequence 1376 BP; 347 A; 338 C; 296 G; 392 T; 3 other;

Query Match 1.5%; Score 17; DB 20; Length 1376;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951  
 |||||  
 Db 1144 agaagttcagaactac 1160

## RESULT 40

T44042  
 ID T44042 standard; cDNA; 1414 BP.

XX  
 AC T44042;

DT 26-FEB-1997 (first entry)

XX Human G-protein chemokine receptor HDGNR10 cDNA.

XX G-protein chemokine receptor; HDGNR10; signal transduction;  
 KW haematopoiesis; leukaemia; inflammation; rheumatoid arthritis;  
 KW diagnosis; gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 259..1317  
 FT /\*tag= a  
 FT primer\_bind complement (259..276)  
 FT /\*tag= b  
 FT /\*note= "5' primer site for cDNA amplification"  
 FT primer\_bind 1297..1314  
 FT /\*tag= c  
 FT /\*note= "3' primer site for cDNA amplification"

XX W09639437-A1.

XX 12-DEC-1996.

XX 06-JUN-1995; 95WO-US07173.

XX 06-JUN-1995; 95WO-US07173.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Li Y, Ruben SM;

XX WPI: 1997-043072/04.

XX P-PSDB; W07602.

XX Human G-protein chemokine receptor, HDGNR10 - useful to identify  
 PT (ant)agonists, for treatment of haematopoiesis, leukaemia, chronic  
 PT and acute inflammation, rheumatoid arthritis, etc.

XX Claim 1: Page 44-46; 61pp; English.

XX A cDNA clone (T44042) codes for the novel human G-protein chemokine

CC receptor HDGNR10 (W07602), a 7-transmembrane protein involved in  
 CC signal transduction. It was discovered in a cDNA library derived  
 CC from human monocytes. The cDNA may be incorporated into a vector  
 CC and utilised in the prodn. of recombinant HDGNR10, as a probe to  
 CC detect mutations in the receptor gene associated with disease, and  
 CC in gene therapy to treat conditions related to underexpression of  
 CC the receptor e.g. to stimulate haematopoiesis, wound healing, or to  
 CC treat solid tumours, autoimmune diseases etc.

XX SQ Sequence 1414 BP; 350 A; 344 C; 313 G; 407 T; 0 other;

Query Match 1.5%; Score 17; DB 18; Length 1414;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951  
 |||||  
 Db 1163 agaagttcagaactac 1179

## RESULT 41

Z91481

ID Z91481 standard; cDNA; 1414 BP.

XX  
 AC Z91481;

XX  
 DT 19-MAY-2000 (first entry)

XX Human G-protein chemokine receptor HDGNR10 encoding cDNA SEQ ID NO:1.

XX Human; G-protein coupled receptor; G-protein chemokine receptor;  
 KW HDGNR10; diagnosis; haematopoiesis; wound healing; coagulation;  
 KW angiogenesis; tumour; infection; leukaemia; psoriasis; allergy;  
 KW T-cell mediated autoimmune disease; atherogenesis; anaphylaxis;  
 KW inflammation; allergic reaction; silicosis; sarcoidosis;  
 KW rheumatoid arthritis; hyper-eosinophilia syndrome; GS.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 259..1317  
 FT /\*tag= a  
 FT /product= "G-protein chemokine receptor HDGNR10"

XX US6025154-A.

XX 15-FEB-2000.

XX 06-JUN-1995; 95US-0466343.

XX 06-JUN-1995; 95US-0466343.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Li Y, Ruben SM;

XX WPI: 2000-181807/16.

XX P-PSDB; Y80128.

XX Isolated nucleic acid encoding human G-protein chemokine receptor  
 PT useful for diagnostic assays, scientific research and screening for  
 PT compounds which bind to and activate or inhibit activation of the  
 PT receptor polypeptides -

XX Claim 9; Fig 1; 22pp; English.

XX The present sequence encodes a human G-protein chemokine receptor  
 CC designated HDGNR10. HDGNR10 polynucleotides are useful in methods of  
 CC screening for compounds which bind to and either: (1) activate the  
 CC HDGNR10 polypeptides causing stimulation of haematopoiesis, wound  
 CC healing, coagulation, and angiogenesis; treatment of solid tumours,  
 CC chronic infections, leukaemia, T-cell mediated autoimmune diseases,

CC parasitic infections, psoriasis, and to stimulate growth factor  
 CC activity; or (2) inhibit activation of the HDGMR10 polypeptides which  
 CC is useful for preventing and/or treating allergy, atherogenesis,  
 CC anaphylaxis, malignancy, chronic and acute inflammation, histamine and  
 CC immunoglobulin E-mediated allergic reactions, prostaglandin-independent  
 CC fever, bone marrow failure, silicosis, sarcoidosis, rheumatoid  
 CC arthritis shock and hyper-eosinophilia syndrome. The polynucleotides  
 CC are also useful for diagnostic assays for detecting diseases related to  
 CC mutations in the nucleic acid sequences encoding the polypeptides and  
 CC for detecting an altered level of the soluble form of the receptor  
 CC polypeptides. The polynucleotides are also useful for in vitro purposes  
 CC related to scientific research, synthesis of DNA and manufacture of  
 CC DNA vectors.

XX  
 SQ Sequence 1414 BP; 350 A; 344 C; 313 G; 407 T; 0 other;

Query Match 1.5%; Score 17; DB 21; Length 1414;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 935 AGAAGTTCAGAACTAC 951  
 Db 1163 agaagttcagaactac 1179  
 |||||

## RESULT 42

T90118  
 ID T90118 standard; cDNA; 1442 BP.

XX  
 AC T90118;

DT 14-APR-1998 (first entry)

XX  
 DE cDNA for inactive human CCR5.

XX Inactive; human Cys-Cys chemokine receptor 5; CCR5;  
 KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;  
 KW predisposition; resistance; diagnosis; treatment; prevention;  
 KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;  
 KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;  
 KW cancer; atherosclerosis; autoimmune disorder; ss.

XX Homo sapiens.

XX  
 FH Key Location/Qualifiers  
 FT CDS 240..887  
 FT /\*tag= a

XX WO9732019-A2.

XX 04-SEP-1997.

XX 28-FEB-1997; 97WO-BE00023.

XX 06-AUG-1996; 96EP-0870102.

XX 01-MAR-1996; 96EP-0870021.

XX (EURO-) EUROSREEN SA.

XX Libert F, Parmentier M, Samson M, Vassart G;

XX WPI; 1997-479829/44.

XX P-PSDB; W27408.

XX Active and inactive forms of human CC chemokine receptor CCR-5 -  
 PT useful to diagnose, prevent and/or treat inflammatory disorders,  
 PT autoimmune disease and viral infection

XX Claim 18; Fig 1d-e; 94pp; English.

XX The present sequence encodes an inactive human CC (Cys-Cys)

CC chemokine receptor 5 (CCR5), which lacks the last 3 transmembrane

CC regions and the regions involved in G protein-coupling. CCR5 or  
 CC its cDNA can be used to diagnose, treat and/or prevent inflammatory  
 CC diseases, e.g. rheumatoid arthritis, glomerulonephritis, asthma,  
 CC idiopathic pulmonary fibrosis and psoriasis, viral infections,  
 CC especially human immunodeficiency virus type 1 or type 2 (HIV-1 or  
 CC HIV-2) infection, cancer, atherosclerosis and autoimmune disorders.  
 CC Subjects that express the inactive receptor have a predisposition,  
 CC or resistance to HIV-1 and/or HIV-2.

XX  
 SQ Sequence 1442 BP; 363 A; 343 C; 314 G; 422 T; 0 other;

Query Match 1.5%; Score 17; DB 18; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 935 AGAAGTTCAGAACTAC 951  
 Db 1112 agaagttcagaactac 1128  
 |||||

## RESULT 43

T90117  
 ID T90117 standard; cDNA; 1477 BP.

XX  
 AC T90117;

DT 14-APR-1998 (first entry)

XX  
 DE cDNA for human CCR5.

XX Human Cys-Cys chemokine receptor 5; CCR5;

KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;

KW diagnosis; treatment; prevention;

KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;

KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;

KW cancer; atherosclerosis; autoimmune disorder; ss.

XX Homo sapiens.

XX  
 FH Key Location/Qualifiers  
 FT CDS 240..1298  
 FT /\*tag= a

XX WO9732019-A2.

XX 04-SEP-1997.

XX 28-FEB-1997; 97WO-BE00023.

XX 06-AUG-1996; 96EP-0870102.

XX 01-MAR-1996; 96EP-0870021.

XX (EURO-) EUROSREEN SA.

XX Libert F, Parmentier M, Samson M, Vassart G;

XX WPI; 1997-479829/44.

XX P-PSDB; T90117.

XX Active and inactive forms of human CC chemokine receptor CCR-5 -  
 PT useful to diagnose, prevent and/or treat inflammatory disorders,  
 PT autoimmune disease and viral infection

XX Claim 18; Fig 1b-c; 94pp; English.

XX The present sequence encodes human CC (Cys-Cys) chemokine receptor  
 CC 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES  
 CC chemokines, but not by monocyte chemoattractant protein 1 (MCP-1),  
 CC MCP-2, MCP-3, interleukin-8 (IL-8) or growth related gene product  
 CC alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of  
 CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).  
 CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent

CC inflammatory diseases, e.g. rheumatoid arthritis,  
 CC glomerulonephritis, asthma, idiopathic pulmonary fibrosis and  
 CC psoriasis, viral infections, especially HIV-1 or HIV-2 infection,  
 CC cancer, atherosclerosis and autoimmune disorders.  
 XX  
 SQ Sequence 1477 BP; 374 A; 349 C; 320 G; 431 T; 3 other;

Query Match 1.5%; Score 17; DB 18; Length 1477;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 935 AGAAGTTCAGAACTAC 951  
 |||||||  
 DB 1144 agaagttcagaactac 1160

RESULT 44  
 A05519  
 ID A05519 standard; DNA; 1521 BP.  
 AC  
 XX  
 AC A05519;  
 XX  
 DT 24-MAY-2000 (first entry)  
 XX  
 DE Streptococcus pneumoniae type 4 nucleotide sequence 4170.4.  
 XX  
 KW Streptococcus pneumoniae; vaccine; screening; protein antigen;  
 KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;  
 KW pneumococcal disease; ds.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN WO200006737-A2.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PF 27-JUL-1999; 99WO-GB02451.  
 XX  
 PR 27-JUL-1999; 98GB-0016337.  
 PR 19-MAR-1999; 99US-0125164.  
 XX  
 PA (MICR-) MICROBIAL TECHNIQS LTD.  
 XX  
 PI Gilbert CFG, Hansbro PM;  
 XX  
 DR WPI; 2000-195300/17.  
 XX  
 PT New Streptococcal protein, useful as a vaccine, for diagnosis of  
 PT pneumococcal diseases and for screening agents capable of antagonizing  
 PT or inhibiting expression of the protein  
 XX  
 PS Claim 7; Page 53; 108pp; English.  
 XX  
 CC Y81501 to Y81679 represent specifically claimed protein sequences  
 CC isolated from Streptococcus pneumoniae. A05407 to A05590 represent  
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.  
 CC The sequences have antibacterial and antiinflammatory properties.  
 CC The protein sequences, and fragments of them, are useful as immunogens  
 CC and/or antigens. The nucleotide sequences can be used in vaccines and in  
 CC diagnostic assays. The proteins and nucleotides can be useful for the  
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also  
 CC useful for screening an agent capable of antagonizing, inhibiting or  
 CC interfering with the function or expression of the proteins in which the  
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection  
 CC and meningitis. A05591 to A05614 represent primers used in the  
 CC exemplification of the present invention.  
 XX  
 SQ Sequence 1521 BP; 407 A; 303 C; 361 G; 450 T; 0 other;

Query Match 1.5%; Score 17; DB 21; Length 1521;  
 Best Local Similarity 100.0%; Pred. No. 75;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 641 TCTTTATCCCTTGATG 657  
 |||||||  
 DB 215 tctttatcccttgatg 231

RESULT 45  
 T80107/c  
 ID T80107 standard; cDNA; 1526 BP.  
 XX  
 AC T80107;  
 XX  
 DT 28-NOV-1997 (first entry)  
 XX  
 DE Human KF-1 gene clone KF361-2.  
 XX  
 KW Alzheimer's disease; KF-1; human; mouse; KF361 gene; antibody; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP09215495-A.  
 XX  
 PD 19-AUG-1997.  
 XX  
 PF 20-JUN-1996; 96JP-0181514.  
 XX  
 PR 07-DEC-1995; 95JP-0345659.  
 XX  
 PA (SUMU ) SUMITOMO SEIYAKU KK.  
 XX  
 DR WPI; 1997-465073/43.  
 XX  
 PT A brain-specific expression gene - used in the diagnosis of  
 PT Alzheimer's disease  
 XX  
 PS Example 3; Page 15-16; 23pp; Japanese.  
 XX  
 CC T80106 and T80107 represent clones of the human KF-1 gene (see T80099 for  
 CC wild type sequence). This sequence, T80098, and T80100 all represent cDNA  
 CC sequences of the invention. These genes, their products and antibodies  
 CC specific for the proteins are useful for the diagnosis of Alzheimer's  
 CC disease.  
 XX  
 SQ Sequence 1526 BP; 436 A; 269 C; 315 G; 506 T; 0 other;

Query Match 1.5%; Score 17; DB 18; Length 1526;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TTTTATAGAGGCCA 205  
 |||||||  
 DB 1031 TTTTATAGAGGCCA 1015

RESULT 46  
 T99542  
 ID T99542 standard; cDNA; 1557 BP.  
 XX  
 AC T99542;  
 XX  
 DT 21-MAY-1998 (first entry)  
 XX  
 DE Human chemokine receptor MMLR-CCR cDNA.  
 XX  
 KW Mammalian mixed lymphocyte receptor; chemokine receptor; MMLR-CCR;  
 KW human; monocyte; macrophage; chemotaxis; haematopoiesis; infection;  
 KW inflammation; proliferative disease; cardiovascular disease;  
 KW tumour; rheumatoid arthritis; alveolitis; atherosclerosis;  
 KW chronic granulomatous disease; asthma; myasthenia gravis;  
 KW diabetes; inflammatory bowel disease; toxic shock syndrome; ss.  
 KW septic shock; Chediak-Higashi syndrome; therapy; diagnosis; ss.

XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 XX FT CDS 36..1034  
 XX FT /\*tag= a  
 XX PN WO9741225-A2.  
 XX PD 06-NOV-1997.  
 XX PF 25-APR-1997; 97WO-US06993.  
 XX PR 26-APR-1996; 96US-0638081.  
 XX PA (INCY-) INCYTE PHARM INC.  
 XX PI Au-Young J, Bandman O, Coleman R, Wilde CG;  
 XX DR WPI; 1997-549729/50.  
 XX DR P-PSDB; W26766.  
 XX PT Polynucleotide encoding MMLR-CCR or MPHG-CCR chemokine receptor -  
 XX PT useful to study, diagnose and treat, e.g. infection, inflammation,  
 XX PT solid tumour and proliferative and cardiovascular disease  
 XX PS Claim 2; Page 36-37; 59pp; English.  
 XX CC This cDNA clone codes for human C-C chemokine receptor MMLR-CCR  
 XX CC (see W26766) that may be associated with monocyte/macrophage  
 XX CC infiltration and chemotaxis and haematopoiesis. It was isolated  
 XX CC from a cDNA library made from mononuclear cells collected on day  
 XX CC 2 of a mixed lymphocyte (MLR) culture, i.e. cells associated with  
 XX CC inflammation and immunomodulation. MMLC-CCR nucleotides are  
 XX CC absent in sequences from a cDNA library of monocytes not subjected  
 XX CC to an MLR. Another clone (see T99543), coding for novel chemokine  
 XX CC receptor MPHG-CCR (see W26767), was derived from monocytes not  
 XX CC subjected to MLR. Also claimed are antisense molecules comprising  
 XX CC the complement of polynucleotides encoding MMLR-CCR or MPHG-CCR,  
 XX CC vectors and transformed host cells. The products can be used to  
 XX CC study, diagnose and treat disease states in which normal leukocyte  
 XX CC function is perturbed by normal leukopoiesis or inappropriate  
 XX CC activation via chemokine agonists or antagonists, such as infection,  
 XX CC inflammation, proliferative disease, tumorigenesis, autoimmune  
 XX CC disease, abnormal cell proliferation, solid tumours, cardiovascular  
 XX CC disease, rheumatoid arthritis, alveolitis, atherosclerosis, chronic  
 XX CC granulomatous disease, asthma, myasthenia gravis, diabetes,  
 XX CC inflammatory bowel disease, toxic shock syndrome, septic shock and  
 XX CC Chediak-Higashi syndrome.  
 XX SQ Sequence 1557 BP; 419 A; 348 C; 327 G; 463 T; 0 other;  
 Query Match 1.5%; Score 17; DB 18; Length 1557;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 935 AGAAGTTCAGAACTAC 951  
 Db 880 agaagttcagaactac 896  
 |||||  
 RESULT 47  
 F21023  
 ID F21023 standard; DNA; 1569 BP.  
 XX AC F21023;  
 XX DT 14-MAR-2001 (first entry)  
 XX DE Human low adenosine antisense oligonucleotide related sequence #2590.  
 XX DE Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
 KW human; airway disorder; bronchoconstriction; lung inflammation;  
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;  
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 KW cancer; ss.  
 XX Homo sapiens.  
 OS WO200062736-A2.  
 XX 26-OCT-2000.  
 XX 24-MAR-2000; 2000WO-US08020.  
 XX 06-APR-1999; 99US-0127958.  
 XX (UYEC-) UNIV EAST CAROLINA.  
 XX (NYCE/) NYCE J W.  
 XX Nyce JW;  
 XX WPI; 2000-679539/66.  
 XX Low adenosine (A) content antisense oligonucleotides which do not  
 XX trigger adenosine receptors during metabolism, useful e.g. for treating  
 XX cancers and respiratory obstructions -  
 XX PS Disclosure; Page 833; 1592pp; English.  
 XX CC The present invention describes low adenosine (A) content antisense  
 XX CC oligonucleotides and compositions (I) comprising them. In the antisense  
 XX CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
 XX CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 XX CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.  
 XX CC The antisense oligonucleotides and (I) can be used to down-regulate the  
 XX CC expression and or activity of target polypeptides associated with  
 XX CC lung/respiratory disorders and malignancies, such as stimulating and  
 XX CC activating peptide factors and transmitters, transcription factors,  
 XX CC immunoglobulins and antibodies, antibody receptors, cytokines and  
 XX CC chemokines, endogenously produced specific and non-specific enzymes,  
 XX CC binding proteins, adhesion molecules and their receptors, cytokine and  
 XX CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
 XX CC nervous system (CNS) and peripheral nervous and non-nervous system  
 XX CC receptors, CNS and peripheral nervous and non-nervous system peptide  
 XX CC transmitters, defensins, growth factors, vasoactive peptides and  
 XX CC receptors, binding proteins and malignancy associated proteins. The  
 XX CC antisense oligonucleotides may be used in this way to treat disorders  
 XX CC including respiratory obstruction (especially pulmonary obstruction  
 XX CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)  
 XX CC and/or surfactant hypoproduction which are associated with a disease or  
 XX CC condition selected from pulmonary vasoconstriction, inflammation,  
 XX CC allergies, asthma, impeded respiration, respiratory distress syndrome  
 XX CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 XX CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 XX CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 XX CC and/or cancer. F18434 to F21543 represent human polynucleotide fragments  
 XX CC and antisense oligonucleotides used in the exemplification of the  
 XX CC present invention.  
 XX SQ Sequence 1569 BP; 440 A; 366 C; 342 G; 421 T; 0 other;  
 Query Match 1.5%; Score 17; DB 21; Length 1569;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 969 GGACCTGGTGTGTCA 985  
 Db 374 ggacctggtgtgtcga 390  
 |||||

Query Match 1.5%; Score 17; DB 21; Length 1569;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 GGACCTGTGTGTGTGA 985  
 DB 374 ggacctgtgtgtgtga 390  
 ||||||||||||||||

RESULT 49  
 T80102/C  
 ID T80102 standard; cDNA; 1641 BP.  
 XX  
 AC T80102;  
 XX  
 DT 25-NOV-1997 (first entry)  
 XX  
 DE KF-1 gene open reading frame.  
 XX  
 KW Alzheimer's disease; KF-1; human; mouse; KF361 gene; antibody; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP09215495-A.  
 XX  
 PD 19-AUG-1997.  
 XX  
 PF 20-JUN-1996; 96JP-0181514.  
 XX  
 PR 07-DEC-1995; 95JP-0345659.  
 XX  
 PA (SUMU) SUMITOMO SEIYAKU KK.  
 XX  
 DR WPI; 1997-465073/43.  
 XX  
 DR P-PSDB; W26164.  
 XX  
 PT A brain-specific expression gene - used in the diagnosis of  
 PT Alzheimer's disease  
 XX  
 PS Claim 6; Page 12; 23pp; Japanese.  
 XX  
 CC This sequence represents the open reading frame of the KF-1 gene.  
 CC The sequences shown in T80098-T80100 all represent cDNAs of the  
 CC invention, and are all KF-1 genes. The genes contain the open reading  
 CC frames represented by this sequence, T80101 and T80103. These genes,  
 CC their products and antibodies specific for the proteins are useful for  
 CC the diagnosis of Alzheimer's disease.  
 XX  
 PS Sequence 1641 BP; 489 A; 306 C; 351 G; 495 T; 0 other;

Query Match 1.5%; Score 17; DB 18; Length 1641;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TTTTATAGAGGCCA 205  
 DB 1589 TTTTATAGAGGCCA 1573  
 ||||||||||||||||

RESULT 50  
 C46304  
 ID C46304 standard; DNA; 1761 BP.  
 XX  
 AC C46304;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 49651.  
 XX  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.

RESULT 48  
 A34901  
 ID A34901 standard; DNA; 1569 BP.  
 XX  
 AC A34901;  
 XX  
 DT 28-JUL-2000 (first entry)  
 XX  
 DE Human adenosine receptor related polynucleotide SEQ ID NO: 2590.  
 XX  
 KW Human; adenosine receptor; low adenosine. antisense oligonucleotide;  
 KW phosphorothioate; impaired respiration; inflammation; allergy;  
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;  
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200009525-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 03-AUG-1999; 99WO-US17712.  
 XX  
 PR 03-AUG-1998; 98US-0095212.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI Nyce JW;  
 XX  
 DR WPI; 2000-205971/18.  
 XX  
 XX New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
 PT cancers  
 XX  
 PS Disclosure; Page 762-763; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the bronchoconstriction and inflammation. A32313 to A35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (A32323 to A33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

Sequence 1569 BP; 440 A; 366 C; 342 G; 421 T; 0 other;

XX	Arabidopsis thaliana.	PR	29-JUN-1999;	99US-0140991.
OS		PR	30-JUN-1999;	99US-0141287.
XX		PR	01-JUL-1999;	99US-0141842.
PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF	25-FEB-2000; 2000EP-0301439.	PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
PR	25-FEB-1999;	PR	13-JUL-1999;	99US-0143542.
PR	05-MAR-1999;	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	28-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999;	PR	21-JUL-1999;	99US-0145088.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070.

PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 17; DB 21; Length 1761;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 CTTTGGCTTTTATAAGA 199  
Db 130 ctttgcctttataaga 146

Search completed: May 24, 2001, 09:25:44  
Job time: 3615 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2001, 07:59:04 ; Search time 1451.51 Seconds  
(without alignments)  
11369.439 Million cell updates/sec

Title: US-08-887-977-9  
Perfect score: 1119  
Sequence: 1 ATGTTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119

Scoring table: OLIGO.NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1283235 seqs, 7373929652 residues

Word size : 15

Total number of hits satisfying chosen parameters: 18528

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl.\*

- 1: gb\_ba1.\*
- 2: gb\_ba2.\*
- 3: gb\_ba3.\*
- 4: gb\_in1.\*
- 5: gb\_in2.\*
- 6: gb\_in3.\*
- 7: gb\_om.\*
- 8: gb\_ov.\*
- 9: gb\_pat1.\*
- 10: gb\_pat2.\*
- 11: gb\_ph.\*
- 12: gb\_pl1.\*
- 13: gb\_pl2.\*
- 14: gb\_pl3.\*
- 15: gb\_pl4.\*
- 16: em\_ba1.\*
- 17: em\_ba2.\*
- 18: em\_fun.\*
- 19: em\_htgo\_hum.\*
- 20: em\_htgo\_inv.\*
- 21: em\_htgo\_rod.\*
- 22: em\_htg\_hum1.\*
- 23: em\_htg\_hum2.\*
- 24: em\_htg\_hum3.\*
- 25: em\_htg\_hum4.\*
- 26: em\_htg\_hum5.\*
- 27: em\_htg\_hum6.\*
- 28: em\_htg\_hum7.\*
- 29: em\_htg\_hum8.\*
- 30: em\_htg\_inv1.\*
- 31: em\_htg\_inv2.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_rod.\*
- 34: em\_hum1.\*
- 35: em\_hum2.\*
- 36: em\_hum3.\*
- 37: em\_hum4.\*
- 38: em\_hum5.\*
- 39: em\_hum6.\*
- 40: em\_hum7.\*
- 41: em\_in.\*
- 42: em\_om.\*
- 43: em\_or.\*

- 44: em\_ov.\*
- 45: em\_pat.\*
- 46: em\_ph.\*
- 47: em\_pl.\*
- 48: em\_ro.\*
- 49: em\_sts.\*
- 50: em\_sy.\*
- 51: em\_un.\*
- 52: em\_vi.\*
- 53: gb\_sts1.\*
- 54: gb\_sts2.\*
- 55: gb\_sts3.\*
- 56: gb\_sy.\*
- 57: gb\_un.\*
- 58: gb\_vil.\*
- 59: gb\_vl2.\*
- 60: gb\_vl2.\*
- 61: gb\_vl2.\*
- 62: gb\_vl2.\*
- 63: gb\_vl2.\*
- 64: gb\_vl2.\*
- 65: gb\_vl2.\*
- 66: gb\_vl2.\*
- 67: gb\_vl2.\*
- 68: gb\_vl2.\*
- 69: gb\_vl2.\*
- 70: gb\_vl2.\*
- 71: gb\_vl2.\*
- 72: gb\_vl2.\*
- 73: gb\_vl2.\*
- 74: gb\_vl2.\*
- 75: gb\_vl2.\*
- 76: gb\_vl2.\*
- 77: gb\_vl2.\*
- 78: gb\_vl2.\*
- 79: gb\_vl2.\*
- 80: gb\_vl2.\*
- 81: gb\_vl2.\*
- 82: gb\_vl2.\*
- 83: gb\_vl2.\*
- 84: gb\_vl2.\*
- 85: gb\_vl2.\*
- 86: gb\_vl2.\*
- 87: gb\_vl2.\*
- 88: gb\_vl2.\*
- 89: gb\_vl2.\*
- 90: gb\_vl2.\*
- 91: gb\_vl2.\*
- 92: gb\_vl2.\*
- 93: gb\_vl2.\*
- 94: gb\_vl2.\*
- 95: gb\_vl2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	869	77.7	1255	92	HSCRL3	Z79784 H.sapiens G
2	869	77.7	3693	93	HSU45984	U45984 Homo sapien
3	818	73.1	101230	92	HSBA517H2	AL121935 Human DNA
4	767	68.5	1518	93	HSU68032	U68032 Human G pro
5	767	68.5	2978	93	HSU68030	U68030 Human G pro
6	442	39.5	1137	93	HSU60000	U60000 Human IL8-r
c 7	434	38.8	205272	79	AL353591	AL353591 Homo sapi
c 8	22	2.0	133588	81	AL390375	AL390375 Homo sapi
c 9	22	2.0	153203	80	AL359812	AL359812 Homo sapi
c 10	22	2.0	161317	81	AL391873	AL391873 Homo sapi
c 11	21	1.9	40020	4	AC068783	AC068783 Leishmani



c 158	18	1.6 113274	80	AL355602	AL355602 Homo sapi	c 231	18	1.6 164232	73	AC067792	AC067792 Homo sapi
c 159	18	1.6 114371	22	AC012017	AC012017 Homo sapi	232	18	1.6 164280	69	AC025391	AC025391 Homo sapi
c 160	18	1.6 114600	91	HS477823	AL033533 Human DNA	c 233	18	1.6 165372	72	AC051613	AC051613 Mus muscu
c 161	18	1.6 115889	81	AL512330	AL512330 Homo sapi	234	18	1.6 165660	71	AC027194	AC027194 Homo sapi
c 162	18	1.6 117069	61	AC008907	AC008907 Homo sapi	235	18	1.6 166105	69	AC024972	AC024972 Homo sapi
c 163	18	1.6 117939	91	HS211A9	Z95889 Human DNA s	236	18	1.6 166179	83	CNS01RH2	AL160313 Homo sapi
c 164	18	1.6 117970	87	AC016651	AC016651 Homo sapi	c 237	18	1.6 166528	73	AC063970	AC063970 Mus muscu
c 165	18	1.6 119563	92	HSJ894D12	AL078605 Human DNA	c 238	18	1.6 167100	80	AL359198	AL359198 Homo sapi
c 166	18	1.6 122562	89	AL136987	AL136987 Human DNA	c 239	18	1.6 168266	62	AC011069	AC011069 Drosophil
c 167	18	1.6 122655	63	AC013189	AC013189 Drosophil	c 240	18	1.6 169848	74	AC069304	AC069304 Homo sapi
c 168	18	1.6 124373	94	AC012104	AC012104 Mus muscu	241	18	1.6 170269	87	AC012085	AC012085 Homo sapi
c 169	18	1.6 128868	92	HSB211023	AL133499 Homo sapi	242	18	1.6 170354	71	AC031992	AC031992 Homo sapi
c 170	18	1.6 129040	83	CNS01DTH	AL132712 Homo sapi	243	18	1.6 170586	83	CNS01RHF	AL161747 Homo sapi
c 171	18	1.6 129625	87	AC009301	AC009301 Homo sapi	244	18	1.6 170673	80	AL359376	AL359376 Homo sapi
c 172	18	1.6 129794	77	AC084027	AC084027 Homo sapi	c 245	18	1.6 170734	72	AC055729	AC055729 Homo sapi
c 173	18	1.6 131546	91	HS14209	AL035687 Human DNA	246	18	1.6 170959	67	AC022197	AC022197 Homo sapi
c 174	18	1.6 132251	61	AC008850	AC008850 Homo sapi	c 247	18	1.6 171916	69	AC025154	AC025154 Homo sapi
c 175	18	1.6 132536	77	AC084393	AC084393 Homo sapi	c 248	18	1.6 171950	91	CNS05TDA	AL356020 Human chr
c 176	18	1.6 134499	12	AC007505	AC007505 Arabidops	249	18	1.6 172474	81	AL451140	AL451140 Homo sapi
c 177	18	1.6 134738	67	AC021488	AC021488 Homo sapi	250	18	1.6 172815	62	AC010920	AC010920 Drosophil
c 178	18	1.6 139270	90	AP001579	AP001579 Homo sapi	c 251	18	1.6 173146	62	AC011243	AC011243 Homo sapi
c 179	18	1.6 140331	86	AC007002	AC007002 Homo sapi	252	18	1.6 173354	71	AC027554	AC027554 Homo sapi
c 180	18	1.6 140680	13	ATFL3112	AL133292 Arabidops	253	18	1.6 173513	91	HS125H2	Z98949 Human DNA s
c 181	18	1.6 141171	61	AC008712	AC008712 Homo sapi	254	18	1.6 175178	62	AC011783	AC011783 Homo sapi
c 182	18	1.6 141944	92	HSJ9061	AL132670 Human DNA	255	18	1.6 175542	81	AL391809	AL391809 Homo sapi
c 183	18	1.6 142008	88	AC025257	AC025257 Homo sapi	256	18	1.6 177076	79	AL157392	AL157392 Homo sapi
c 184	18	1.6 145124	68	AC022894	AC022894 Homo sapi	c 257	18	1.6 177516	75	AC072056	AC072056 Homo sapi
c 185	18	1.6 145353	71	AC034177	AC034177 Homo sapi	c 258	18	1.6 177549	22	AC012012	AC012012 Homo sapi
c 186	18	1.6 145888	75	AC073219	AC073219 Homo sapi	c 259	18	1.6 177627	69	AC024404	AC024404 Homo sapi
c 187	18	1.6 145942	77	AC080001	AC080001 Homo sapi	c 260	18	1.6 177627	75	AC073906	AC073906 Trypanoso
c 188	18	1.6 146153	4	AC007452	AC007452 Drosophil	c 261	18	1.6 177770	72	AC051649	AC051649 Homo sapi
c 189	18	1.6 146637	67	AC021597	AC021597 Homo sapi	c 262	18	1.6 177891	70	AC026810	AC026810 Homo sapi
c 190	18	1.6 146927	75	AC074032	AC074032 Homo sapi	263	18	1.6 178252	80	AL356386	AL356386 Homo sapi
c 191	18	1.6 147054	80	AL356582	AL356582 Homo sapi	264	18	1.6 178422	72	AC051644	AC051644 Homo sapi
c 192	18	1.6 148209	72	AC036178	AC036178 Homo sapi	265	18	1.6 179041	73	AC064802	AC064802 Homo sapi
c 193	18	1.6 148334	77	AC087150	AC087150 Homo sapi	266	18	1.6 179061	78	AL139007	AL139007 Homo sapi
c 194	18	1.6 149848	79	AL162590	AL162590 Homo sapi	c 267	18	1.6 179376	68	AC022809	AC022809 Homo sapi
c 195	18	1.6 151892	69	AC025144	AC025144 Homo sapi	c 268	18	1.6 180596	65	AC019023	AC019023 Homo sapi
c 196	18	1.6 152434	64	AC016262	AC016262 Homo sapi	c 269	18	1.6 181245	79	AL354815	AL354815 Homo sapi
c 197	18	1.6 152464	71	AC027421	AC027421 Homo sapi	c 270	18	1.6 181283	90	AP000567	AP000567 Homo sapi
c 198	18	1.6 152782	74	AC069484	AC069484 Homo sapi	271	18	1.6 181576	88	AC084732	AC084732 Homo sapi
c 199	18	1.6 153041	78	AC087316	AC087316 Homo sapi	272	18	1.6 181819	74	AC069381	AC069381 Homo sapi
c 200	18	1.6 153863	69	AC025361	AC025361 Homo sapi	273	18	1.6 182047	63	AC013462	AC013462 Homo sapi
c 201	18	1.6 153882	61	AC009172	AC009172 Homo sapi	274	18	1.6 182467	63	AC012203	AC012203 Homo sapi
c 202	18	1.6 154711	68	AC023273	AC023273 Homo sapi	c 275	18	1.6 182959	83	CNS01RHW	AL162551 Homo sapi
c 203	18	1.6 155140	74	AC068374	AC068374 Homo sapi	c 276	18	1.6 183345	74	AC068519	AC068519 Homo sapi
c 204	18	1.6 155510	79	AL353723	AL353723 Homo sapi	c 277	18	1.6 183862	78	AC087245	AC087245 Homo sapi
c 205	18	1.6 156316	79	AL157946	AL157946 Homo sapi	278	18	1.6 185091	64	AC015835	AC015835 Homo sapi
c 206	18	1.6 156515	68	AC022947	AC022947 Homo sapi	c 279	18	1.6 185091	64	AC015835	AC015835 Homo sapi
c 207	18	1.6 156614	63	AC013366	AC013366 Homo sapi	280	18	1.6 185823	70	AC026356	AC026356 Homo sapi
c 208	18	1.6 156861	70	AC025739	AC025739 Homo sapi	281	18	1.6 185864	60	AC007598	AC007598 Homo sapi
c 209	18	1.6 157100	80	AL356263	AL356263 Homo sapi	c 282	18	1.6 186062	70	AC026048	AC026048 Homo sapi
c 210	18	1.6 157185	73	AC064798	AC064798 Homo sapi	c 283	18	1.6 186063	68	AC023558	AC023558 Homo sapi
c 211	18	1.6 157201	67	AC021820	AC021820 Homo sapi	284	18	1.6 186137	69	AC024224	AC024224 Homo sapi
c 212	18	1.6 157986	70	AC025860	AC025860 Homo sapi	285	18	1.6 187166	77	AC084403	AC084403 Mus muscu
c 213	18	1.6 158483	70	AC025987	AC025987 Homo sapi	c 286	18	1.6 187593	73	AC067950	AC067950 Homo sapi
c 214	18	1.6 159155	67	AC022293	AC022293 Homo sapi	287	18	1.6 187999	80	AL359081	AL359081 Homo sapi
c 215	18	1.6 159637	60	AC007990	AC007990 Homo sapi	288	18	1.6 188818	65	AC017074	AC017074 Homo sapi
c 216	18	1.6 160130	83	AP003049	AP003049 Homo sapi	c 289	18	1.6 189573	68	AC023842	AC023842 Homo sapi
c 217	18	1.6 160750	92	HSJ177110	AL080284 Human DNA	290	18	1.6 189736	61	AC009563	AC009563 Homo sapi
c 218	18	1.6 160771	91	HS88L2	AL035091 Human DNA	c 291	18	1.6 190000	60	AC004479	AC004479 Homo sapi
c 219	18	1.6 161254	62	AC011829	AC011829 Homo sapi	292	18	1.6 190317	77	AC084332	AC084332 Homo sapi
c 220	18	1.6 161453	85	AC004993	AC004993 Homo sapi	c 293	18	1.6 190429	71	AC026916	AC026916 Homo sapi
c 221	18	1.6 161908	62	AC011989	AC011989 Homo sapi	294	18	1.6 190705	66	AC019331	AC019331 Homo sapi
c 222	18	1.6 161993	79	AL353622	AL353622 Homo sapi	295	18	1.6 191631	80	AL356266	AL356266 Homo sapi
c 223	18	1.6 162296	66	AC020702	AC020702 Homo sapi	296	18	1.6 192800	76	AC074289	AC074289 Homo sapi
c 224	18	1.6 162394	80	AL356773	AL356773 Homo sapi	297	18	1.6 193567	61	AC008945	AC008945 Homo sapi
c 225	18	1.6 162384	60	AC007840	AC007840 Homo sapi	298	18	1.6 193999	72	AC048361	AC048361 Mus muscu
c 226	18	1.6 163396	89	AL161445	AL161445 Human DNA	299	18	1.6 194135	65	AC017081	AC017081 Homo sapi
c 227	18	1.6 163446	75	AC073321	AC073321 Homo sapi	300	18	1.6 194635	85	AC012087	AC012087 Homo sapi
c 228	18	1.6 163573	70	AC026780	AC026780 Homo sapi	c 301	18	1.6 195347	67	AC018877	AC018877 Homo sapi
c 229	18	1.6 164013	90	AL360088	AL360088 Human DNA	c 302	18	1.6 198918	71	AC032022	AC032022 Homo sapi
c 230	18	1.6 164039	74	AC069354	AC069354 Homo sapi	c 303	18	1.6 200327	67	AC021532	AC021532 Homo sapi

304.	18	1.6	200349	93	HSU85195	U85195 Homo sapien	377	17	1.5	1019	88	AF161898	Pan trogl
305	18	1.6	200727	63	AC012296	AL13296 Mus muscu	378	17	1.5	1019	88	AF161899	Pan trogl
c 306	18	1.6	201410	92	HSP12	AL13500 Homo sapi	379	17	1.5	1019	88	AF161900	Pan trogl
c 307	18	1.6	202689	71	AC034281	AC034281 Homo sapi	380	17	1.5	1019	88	AF161901	Pan trogl
c 308	18	1.6	206210	67	AC021648	AC021648 Homo sapi	381	17	1.5	1019	88	AF161902	Pan trogl
c 309	18	1.6	206979	78	AC087541	AC087541 Mus muscu	382	17	1.5	1019	88	AF161903	Pan trogl
c 310	18	1.6	207862	63	AC012421	AC012421 Homo sapi	383	17	1.5	1019	88	AF161904	Pan trogl
311	18	1.6	211533	74	AC068807	AC068807 Mus muscu	384	17	1.5	1019	88	AF161905	Pan trogl
312	18	1.6	212390	80	AL356275	AL356275 Homo sapi	385	17	1.5	1019	88	AF161906	Pongo pyg
c 313	18	1.6	212835	60	AC007612	AC007612 Homo sapi	386	17	1.5	1019	88	AF161907	Pongo pyg
c 314	18	1.6	215158	61	AC009098	AC009098 Homo sapi	387	17	1.5	1019	88	AF161908	Pongo pyg
c 315	18	1.6	220480	68	AC023973	AC023973 Homo sapi	388	17	1.5	1019	88	AF161909	Homo sapi
c 316	18	1.6	221003	83	CNS01DRW	AL121576 Homo sapi	389	17	1.5	1019	88	AF161910	Homo sapi
c 317	18	1.6	224474	72	AC048380	AC048380 Homo sapi	390	17	1.5	1019	88	AF161911	Homo sapi
c 318	18	1.6	235030	60	AC007601	AC007601 Homo sapi	391	17	1.5	1019	88	AF161912	Homo sapi
c 319	18	1.6	240864	60	AC006510	AC006510 Homo sapi	392	17	1.5	1019	88	AF161913	Homo sapi
320	18	1.6	245708	88	AF250324	AF250324 Homo sapi	393	17	1.5	1019	88	AF161914	Homo sapi
321	18	1.6	250529	93	HUA000658	AE000658 Homo sapi	394	17	1.5	1019	88	AF161915	Homo sapi
c 322	18	1.6	251664	61	AC008908	AC008908 Homo sapi	395	17	1.5	1019	88	AF161916	Homo sapi
c 323	18	1.6	261002	83	CXY41C4	AL022278 Caenorhab	396	17	1.5	1019	88	AF161917	Homo sapi
c 324	18	1.6	261002	83	CXY41C4	AL022278 Caenorhab	397	17	1.5	1019	88	AF161919	Homo sapi
325	18	1.6	264646	4	AE003818	AE003818 Drosophila	398	17	1.5	1019	88	AF161920	Homo sapi
326	18	1.6	280278	81	AL390201	AL390201 Homo sapi	399	17	1.5	1019	88	AF161921	Homo sapi
327	18	1.6	280810	72	AC046188	AC046188 Homo sapi	400	17	1.5	1019	88	AF161948	Cercopith
c 328	18	1.6	286583	4	AE003565	AE003565 Drosophila	401	17	1.5	1019	88	AF161948	Cercopith
c 329	18	1.6	303209	4	AE003604	AE003604 Drosophila	402	17	1.5	1019	88	AF161950	Macaca fa
330	18	1.6	303356	4	AE003583	AE003583 Drosophila	403	17	1.5	1019	88	AF161951	Macaca fa
c 331	18	1.6	308092	4	AE003493	AE003493 Drosophila	404	17	1.5	1019	88	AF161952	Macaca fa
c 332	18	1.6	340000	90	AP001708	AP001708 Homo sapi	405	17	1.5	1019	88	AF161953	Macaca fa
c 333	18	1.6	340000	91	HS21C101	AL163301 Homo sapi	406	17	1.5	1019	88	AF161954	Macaca fa
c 334	18	1.6	347550	2	AP001118	AP001118 Buchnera	407	17	1.5	1019	88	AF161955	Macaca fu
c 335	17	1.5	67	2	ECOLALPHON	M37336 Escherichia	408	17	1.5	1019	88	AF161956	Macaca fu
336	17	1.5	172	93	HUMSPEC23	M61796 Human alpha	409	17	1.5	1019	88	AF161957	Macaca fu
c 337	17	1.5	197	88	AF162701	AF162701 Homo sapi	410	17	1.5	1019	88	AF161959	Macaca mu
338	17	1.5	340	54	G35559	G35559 STS h14a541	411	17	1.5	1019	88	AF161960	Macaca mu
339	17	1.5	343	88	AH006349S4	M76663 Human h11-be	412	17	1.5	1019	88	AF161961	Macaca mu
c 340	17	1.5	349	54	G02525	G02525 human STS W	413	17	1.5	1019	88	AF161962	Macaca mu
c 341	17	1.5	355	54	BTMSAT85	X71578 B. taurus mt	414	17	1.5	1019	88	AF161964	Macaca mu
c 342	17	1.5	400	94	MMIRS	X61175 Mouse MHC H	415	17	1.5	1019	88	AF161965	Macaca mu
c 343	17	1.5	455	54	G34387	G34387 human STS S	416	17	1.5	1019	88	AF161966	Macaca mu
c 344	17	1.5	480	92	HSANKONE39	U50130 Homo sapien	417	17	1.5	1019	88	AF161967	Macaca mu
c 345	17	1.5	487	54	G62500	G62500 SHGC-140077	418	17	1.5	1019	88	AF161968	Macaca mu
346	17	1.5	500	1	AB038675	AB038675 Unculture	419	17	1.5	1019	88	AF161969	Macaca mu
347	17	1.5	500	1	AB046933	AB046933 Unculture	420	17	1.5	1019	88	AF161970	Macaca mu
348	17	1.5	554	54	G60950	G60950 SHGC-84348	421	17	1.5	1019	88	AF161971	Macaca mu
349	17	1.5	572	93	HST000258	AL365522 Homo sapi	422	17	1.5	1019	88	AF161972	Macaca mu
c 350	17	1.5	596	94	MUSMHLDC	M12483 Mouse MHC c	423	17	1.5	1019	88	AF161973	Macaca mu
c 351	17	1.5	683	14	PTA005758	AJ005758 Populus t	424	17	1.5	1019	88	AF161974	Macaca mu
c 352	17	1.5	683	88	AF229547	AF229547 Homo sapi	425	17	1.5	1019	88	AF161975	Macaca mu
c 353	17	1.5	686	13	ATHMGBET1	Y14072 Arabidopsis	426	17	1.5	1019	88	AF161976	Macaca ne
c 354	17	1.5	711	54	G58833	G58833 SHGC-105905	427	17	1.5	1019	88	AF161977	Macaca ne
c 355	17	1.5	731	13	ATHMGBET2	Y14073 Arabidopsis	428	17	1.5	1019	88	AF161978	Macaca ne
c 356	17	1.5	734	15	STHGHMOB	AJ002391 Solanum t	429	17	1.5	1019	88	AF161979	Macaca ne
c 357	17	1.5	738	13	ATY09482	Y09482 A. thaliana	430	17	1.5	1019	88	AF161980	Macaca ne
358	17	1.5	750	93	HST000256	AL365520 Homo sapi	431	17	1.5	1019	88	AF161981	Macaca ne
359	17	1.5	897	6	TRVSGRS	K01633 T. brucei ta	432	17	1.5	1019	88	AF161982	Macaca ne
360	17	1.5	898	53	CNS06MUX	AL405967 T3 end of	433	17	1.5	1019	88	AF161984	Macaca ne
c 361	17	1.5	938	2	ECORLPB	M18277 E.coli rlpB	434	17	1.5	1019	88	AF161985	Macaca ne
362	17	1.5	1003	58	AF162049	AF162049 Erythroce	435	17	1.5	1019	88	AF161986	Macaca ne
363	17	1.5	1007	53	CNS06C2T	AL393183 T3 end of	436	17	1.5	1019	88	AF161987	Papio cyn
364	17	1.5	1007	53	CNS06MDQ	AL405348 T3 end of	437	17	1.5	1019	88	AF161988	Papio cyn
365	17	1.5	1018	88	AF161918	AF161918 Homo sapi	438	17	1.5	1019	88	AF161989	Papio cyn
366	17	1.5	1019	88	AF161887	AF161887 Hylobates	439	17	1.5	1019	88	AF161990	Papio cyn
367	17	1.5	1019	88	AF161888	AF161888 Hylobates	440	17	1.5	1019	88	AF161991	Papio cyn
368	17	1.5	1019	88	AF161889	AF161889 Hylobates	441	17	1.5	1019	88	AF161992	Papio cyn
369	17	1.5	1019	88	AF161890	AF161890 Hylobates	442	17	1.5	1019	88	AF161994	Papio cyn
370	17	1.5	1019	88	AF161891	AF161891 Gorilla g	443	17	1.5	1019	88	AF161996	Papio cyn
371	17	1.5	1019	88	AF161892	AF161892 Gorilla g	444	17	1.5	1019	88	AF161997	Papio cyn
372	17	1.5	1019	88	AF161893	AF161893 Gorilla g	445	17	1.5	1019	88	AF161998	Mandrillu
373	17	1.5	1019	88	AF161894	AF161894 Gorilla g	446	17	1.5	1019	88	AF161999	Mandrillu
374	17	1.5	1019	88	AF161895	AF161895 Gorilla g	447	17	1.5	1019	88	AF162000	Colobus g
375	17	1.5	1019	88	AF161896	AF161896 Gorilla g	448	17	1.5	1019	88	AF162001	Colobus g
376	17	1.5	1019	88	AF161897	AF161897 Gorilla g	449	17	1.5	1019	88	AF162002	Colobus g

450	17	1.5	1019	88	AF162003	AF162003 Colobus g	523	17	1.5	1059	88	AF035223	Cercopith
451	17	1.5	1019	88	AF162004	AF162004 Colobus g	524	17	1.5	1059	88	AF051902	Cercocebu
452	17	1.5	1019	88	AF162005	AF162005 Colobus g	525	17	1.5	1059	88	AF051903	Cercocebu
453	17	1.5	1019	88	AF162006	AF162006 Chloroceb	526	17	1.5	1059	88	AF051904	Cercocebu
454	17	1.5	1019	88	AF162007	AF162007 Chloroceb	527	17	1.5	1059	88	AF051905	Cercocebu
455	17	1.5	1019	88	AF162015	AF162015 Saguinus	528	17	1.5	1059	88	AF052539	Homo sapi
456	17	1.5	1019	88	AF162016	AF162016 Chloroceb	529	17	1.5	1059	88	AF075442	Trachypit
457	17	1.5	1019	88	AF162017	AF162017 Chloroceb	530	17	1.5	1059	88	AF075443	Trachypit
458	17	1.5	1019	88	AF162018	AF162018 Chloroceb	531	17	1.5	1059	88	AF075444	Pygathrix
459	17	1.5	1019	88	AF162019	AF162019 Chloroceb	532	17	1.5	1059	88	AF075445	Pygathrix
460	17	1.5	1019	88	AF162020	AF162020 Cercopith	533	17	1.5	1059	88	AF075446	Pongo pyg
461	17	1.5	1019	88	AF162022	AF162022 Chloroceb	534	17	1.5	1059	88	AF075447	Pygathrix
462	17	1.5	1019	88	AF162023	AF162023 Chloroceb	535	17	1.5	1059	88	AF075448	Pygathrix
463	17	1.5	1019	88	AF162024	AF162024 Hyllobates	536	17	1.5	1059	88	AF075449	Macaca ar
464	17	1.5	1019	88	AF162026	AF162026 Cercopith	537	17	1.5	1059	88	AF075450	Macaca ar
465	17	1.5	1019	88	AF162027	AF162027 Chloroceb	538	17	1.5	1059	88	AF075451	Hyllobates
466	17	1.5	1019	88	AF162029	AF162029 Chloroceb	539	17	1.5	1059	88	AF079472	Cercocebu
467	17	1.5	1019	88	AF162030	AF162030 Chloroceb	540	17	1.5	1059	88	AF081577	Cercopith
468	17	1.5	1019	88	AF162031	AF162031 Chloroceb	541	17	1.5	1059	88	AF081578	Cercopith
469	17	1.5	1019	88	AF162032	AF162032 Mandrillu	542	17	1.5	1059	88	AF081579	Cercopith
470	17	1.5	1019	88	AF162033	AF162033 Erythroce	543	17	1.5	1059	88	AF084004	Cercocebu
471	17	1.5	1019	88	AF162034	AF162034 Erythroce	544	17	1.5	1059	88	AF105281	Macaca ne
472	17	1.5	1019	88	AF162035	AF162035 Erythroce	545	17	1.5	1059	88	AF105282	Macaca ne
473	17	1.5	1019	88	AF162036	AF162036 Erythroce	546	17	1.5	1059	88	AF105283	Macaca ar
474	17	1.5	1019	88	AF162037	AF162037 Cercopith	547	17	1.5	1059	88	AF105284	Chloroceb
475	17	1.5	1019	88	AF162038	AF162038 Cercopith	548	17	1.5	1059	88	AF105285	Chloroceb
476	17	1.5	1019	88	AF162039	AF162039 Cercopith	549	17	1.5	1059	88	AF105286	Chloroceb
477	17	1.5	1019	88	AF162040	AF162040 Cercopith	550	17	1.5	1059	88	AF105287	Papio ham
478	17	1.5	1019	88	AF162041	AF162041 Cercopith	551	17	1.5	1059	88	AF105288	Papio ham
479	17	1.5	1019	88	AF162042	AF162042 Cercopith	552	17	1.5	1059	88	AF105289	Papio ham
480	17	1.5	1019	88	AF162043	AF162043 Cercopith	553	17	1.5	1059	88	AF105290	Papio ham
481	17	1.5	1019	88	AF162044	AF162044 Cercopith	554	17	1.5	1059	88	AF105291	Gorilla g
482	17	1.5	1019	88	AF162045	AF162045 Cercopith	555	17	1.5	1059	88	AF141639	Colobus g
483	17	1.5	1019	88	AF162046	AF162046 Cercopith	556	17	1.5	1059	88	AF141640	Papio ham
484	17	1.5	1019	88	AF162047	AF162047 Cercopith	557	17	1.5	1059	88	AF141641	Cercopith
485	17	1.5	1019	88	AF162048	AF162048 Erythroce	558	17	1.5	1059	88	AF212100	Cercopith
486	17	1.5	1019	88	AF162050	AF162050 Erythroce	559	17	1.5	1059	88	AF212101	Cercopith
487	17	1.5	1019	88	AF162051	AF162051 Erythroce	560	17	1.5	1059	88	AF212102	Cercopith
488	17	1.5	1019	88	AF162052	AF162052 Erythroce	561	17	1.5	1059	88	AF252551	Hyllobates
489	17	1.5	1019	88	AF162053	AF162053 Erythroce	562	17	1.5	1059	88	AF252552	Chloroceb
490	17	1.5	1019	88	AF162054	AF162054 Erythroce	563	17	1.5	1059	93	MMU73739	U73739 Macaca mula
491	17	1.5	1019	88	AF162055	AF162055 Erythroce	564	17	1.5	1059	93	MMU96762	U96762 Macaca mula
492	17	1.5	1020	88	AF161958	AF161958 Macaca mu	565	17	1.5	1059	93	PTU94329	PTU94329 Pan troglod
493	17	1.5	1020	88	AF161963	AF161963 Macaca mu	566	17	1.5	1060	8	AST251658	AST251658 Acipenser
494	17	1.5	1027	92	HSCMBKR5	X99393 H. sapiens C	567	17	1.5	1060	88	AF011500	Homo sapi
495	17	1.5	1027	93	HSCMBKR5	U66285 Human CC ch	568	17	1.5	1060	88	AF011501	Homo sapi
496	17	1.5	1032	10	I65512	I65512 Sequence 6	569	17	1.5	1060	88	AF011502	Homo sapi
497	17	1.5	1035	88	AF079473	AF079473 Cercocebu	570	17	1.5	1060	88	AF011503	Homo sapi
498	17	1.5	1035	88	AF084002	AF084002 Cercocebu	571	17	1.5	1060	88	AF011505	Homo sapi
499	17	1.5	1035	88	AF084003	AF084003 Cercocebu	572	17	1.5	1060	88	AF011506	Homo sapi
500	17	1.5	1035	88	AF094752	AF094752 Cercocebu	573	17	1.5	1060	88	AF011507	Homo sapi
501	17	1.5	1035	88	AF094753	AF094753 Cercocebu	574	17	1.5	1060	88	AF011508	Homo sapi
502	17	1.5	1037	8	GGYRM2	X04353 Chicken bur	575	17	1.5	1060	88	AF011509	Homo sapi
503	17	1.5	1058	7	AF238859	AF238859 Bos tauru	576	17	1.5	1060	88	AF011510	Homo sapi
504	17	1.5	1059	7	AF202997	AF202997 Bos tauru	577	17	1.5	1060	88	AF011511	Homo sapi
505	17	1.5	1059	85	AB015944	AB015944 Chloroceb	578	17	1.5	1060	88	AF011512	Homo sapi
506	17	1.5	1059	88	AF005658	AF005658 Papio ham	579	17	1.5	1060	88	AF011514	Homo sapi
507	17	1.5	1059	88	AF005659	AF005659 Gorilla g	580	17	1.5	1060	88	AF011516	Homo sapi
508	17	1.5	1059	88	AF005860	AF005860 Macaca fa	581	17	1.5	1060	88	AF011517	Homo sapi
509	17	1.5	1059	88	AF005661	AF005661 Macaca ne	582	17	1.5	1060	88	AF011518	Homo sapi
510	17	1.5	1059	88	AF005662	AF005662 Macaca mu	583	17	1.5	1060	88	AF011519	Homo sapi
511	17	1.5	1059	88	AF005663	AF005663 Pan trogl	584	17	1.5	1060	88	AF011521	Homo sapi
512	17	1.5	1059	88	AF011504	AF011504 Homo sapi	585	17	1.5	1060	88	AF011522	Homo sapi
513	17	1.5	1059	88	AF023452	AF023452 Papio ham	586	17	1.5	1060	88	AF011523	Homo sapi
514	17	1.5	1059	88	AF035214	AF035214 Pan trogl	587	17	1.5	1060	88	AF011525	Homo sapi
515	17	1.5	1059	88	AF035215	AF035215 Cercocebu	588	17	1.5	1060	88	AF011526	Homo sapi
516	17	1.5	1059	88	AF035216	AF035216 Cercopith	589	17	1.5	1060	88	AF011527	Homo sapi
517	17	1.5	1059	88	AF035217	AF035217 Cercopith	590	17	1.5	1060	88	AF011528	Homo sapi
518	17	1.5	1059	88	AF035218	AF035218 Cercopith	591	17	1.5	1060	88	AF011529	Homo sapi
519	17	1.5	1059	88	AF035219	AF035219 Cercopith	592	17	1.5	1060	88	AF011530	Homo sapi
520	17	1.5	1059	88	AF035220	AF035220 Cercopith	593	17	1.5	1060	88	AF011531	Homo sapi
521	17	1.5	1059	88	AF035221	AF035221 Cercopith	594	17	1.5	1060	88	AF011532	Homo sapi
522	17	1.5	1059	88	AF035222	AF035222 Cercopith	595	17	1.5	1060	88	AF011533	Homo sapi

596	17	1.5	1060	88	AF0111534	AF0111534 Homo sapi	669	17	1.5	3383	93	HSU54994	U54994 Human CC ch
597	17	1.5	1060	88	AF0111535	AF0111535 Homo sapi	c 670	17	1.5	3423	10	E13590	E13590 cDNA encodi
598	17	1.5	1060	88	AF0111536	AF0111536 Homo sapi	c 671	17	1.5	3423	91	D76444	D76444 Homo sapien
599	17	1.5	1060	88	AF0111537	AF0111537 Homo sapi	c 672	17	1.5	3526	10	E13586	E13586 cDNA encodi
600	17	1.5	1060	88	AF0111538	AF0111538 Pan trogl	c 673	17	1.5	3573	3	SCY13833	Y13833 Streptomyce
601	17	1.5	1060	88	AF0111539	AF0111539 Pan trogl	c 674	17	1.5	4107	94	MMU20238	MU20238 Mus musculu
602	17	1.5	1060	88	AF0111540	AF0111540 Pan trogl	c 675	17	1.5	4731	94	MUSMHHK2	MUSMHHK2 Mouse MHC c
603	17	1.5	1060	88	AF0111541	AF0111541 Pan trogl	c 676	17	1.5	4849	15	YLU65409	Y65409 Yarrowia li
604	17	1.5	1060	88	AF0111542	AF0111542 Pan trogl	c 677	17	1.5	4855	94	ABU26817	ABU26817 Mus muscu
605	17	1.5	1075	88	AF0119379	AF0119379 Cercopith	c 678	17	1.5	5221	94	MUSMHC04A	M18837 Mouse MHC c
606	17	1.5	1084	93	PTU89797	PTU89797 Pan troglod	c 679	17	1.5	5479	94	MMH2KK	X01652 Mouse H-2K(
607	17	1.5	1086	90	CAU83324	CAU83324 Cercopithec	c 680	17	1.5	5993	94	MUSMHTA	M14825 Mouse MHC c
608	17	1.5	1086	90	CAU83325	CAU83325 Cercopithec	c 681	17	1.5	6036	92	HSBCL7AIG	X90000 H.sapiens D
609	17	1.5	1089	94	HSU83326	HSU83326 Human CC ch	c 682	17	1.5	6059	92	HSBCL7AIG	AF031237 Homo sapi
610	17	1.5	1089	13	PTRECLF	L34815 Pyramimonas	c 683	17	1.5	6192	93	HUMANK	M28880 Human eryth
611	17	1.5	1092	93	U97666	U97666 Pan troglod	c 684	17	1.5	6384	85	ABO33069	ABO33069 Homo sapi
c 612	17	1.5	1147	2	ECODNAPS	M94267 E. coli DNA	c 685	17	1.5	6581	8	SSGGH	X61938 S.salar gen
c 613	17	1.5	1154	53	CNSO6RYF	AL412573 T7 end of	c 686	17	1.5	6839	94	AF001797	AF001797 Mus muscu
c 614	17	1.5	1219	93	HUMIGHWC	M35529 Human Ig he	c 687	17	1.5	7137	11	ABO20525	ABO20525 Bacteriop
c 615	17	1.5	1220	93	HST814X	X12502 H.sapiens t	c 688	17	1.5	7252	92	HSANKY	X16609 Human mRNA
c 616	17	1.5	1225	93	HSU57840	U57840 Human CC ch	c 689	17	1.5	7358	5	AF117240	AF117240 Trypanoso
c 617	17	1.5	1261	2	ECODELTA	L04576 Escherichia	c 690	17	1.5	7422	88	AF009962	AF009962 Homo sapi
c 618	17	1.5	1278	13	AF270844	AF270844 Saccharom	c 691	17	1.5	8001	93	HUMSP1A01	M61877 Human eryth
c 619	17	1.5	1284	8	AF021141	AF021141 Gallus ga	c 692	17	1.5	9224	6	DMALD	X60064 D.melanogas
c 620	17	1.5	1286	94	MUSUTROPH	L22094 Mouse utrop	c 693	17	1.5	10001	1	AE002248	AE002248 Chlamydo
c 621	17	1.5	1310	1	AF059036	AF059036 Pasteurel	c 694	17	1.5	10407	1	AE005243	AE005243 Escherich
c 622	17	1.5	1314	7	AF061521	AF061521 Bos tauru	c 695	17	1.5	11240	1	AE000864	AE000864 Methanoba
c 623	17	1.5	1327	85	AB041391	AB041391 Homo sapi	c 696	17	1.5	11781	1	AE001676	AE001676 Moraxella
c 624	17	1.5	1327	85	AB041393	AB041393 Gorilla g	c 697	17	1.5	12082	1	AF129811	AF129811 Moraxella
c 625	17	1.5	1332	85	AB041392	AB041392 Pan trogl	c 698	17	1.5	12663	1	AE000168	AE000168 Escherich
c 626	17	1.5	1376	92	HSCCKRAG	X91492 H.sapiens C	c 699	17	1.5	12948	1	AE006162	AE006162 Pasteurel
c 627	17	1.5	1439	94	MMILGFBI	X81579 M.musculus	c 700	17	1.5	12994	66	AC019946	AC019946 Drosophil
c 628	17	1.5	1503	6	DROFI6BP01	M98352 Drosophila	c 701	17	1.5	14160	1	AE000604	AE000604 Helicobac
c 629	17	1.5	1521	1	AF162656	AF162656 Streptoco	c 702	17	1.5	14401	1	AE001518	AE001518 Helicobac
c 630	17	1.5	1526	10	E13587	M22329 Rat cytochr	c 703	17	1.5	16496	64	AC014316	AC014316 Drosophil
c 631	17	1.5	1542	94	RATCYPB	M23239 Rat cytochr	c 704	17	1.5	16996	92	HSL21F12B	Z68885 Human DNA s
c 632	17	1.5	1550	93	HUMIGJGEN	M25625 Human Igd a	c 705	17	1.5	17243	92	HSL241B9C	Z69708 Human DNA s
c 633	17	1.5	1569	92	HSWM2	X15264 Human gene	c 706	17	1.5	17641	1	AE001804	AE001804 Escherichia
c 634	17	1.5	1589	2	BACGPRA	M55262 Bacillus me	c 707	17	1.5	18657	2	D90704	D90704 Thermotog
c 635	17	1.5	1590	94	RATCY2DC	J02869 Rat cytochr	c 708	17	1.5	21349	64	AC014598	AC014598 Drosophil
c 636	17	1.5	1735	94	AF047727	AF047727 Mus muscu	c 709	17	1.5	22209	94	AF100154	AF100154 Rattus no
c 637	17	1.5	1747	8	AF000239	AF000239 Xenopus l	c 710	17	1.5	22516	5	CEC50C10	Z72505 Caenorhabdi
c 638	17	1.5	1770	94	AB021743	AB021743 Mus muscu	c 711	17	1.5	23703	94	AF220365	AF220365 Mus muscu
c 639	17	1.5	1855	8	AF169794	AF169794 Xenopus l	c 712	17	1.5	24038	89	AL133277	AL133277 Human DNA
c 640	17	1.5	1966	93	MMU77672	U77672 Macaca mula	c 713	17	1.5	24688	5	CEFI5C11	Z71260 Caenorhabdi
c 641	17	1.5	2001	92	HSPMEL03	U31797 Human melan	c 714	17	1.5	25193	88	AF200923	AF200923 Homo sapi
c 642	17	1.5	2028	94	MSH2KBM1	X15104 Mouse H-2K(	c 715	17	1.5	26329	78	AC087354	AC087354 Homo sapi
c 643	17	1.5	2049	94	MMH2PRO	X54888 Murine H-2K	c 716	17	1.5	26674	93	HSU116E7	Z70273 Human DNA s
c 644	17	1.5	2056	8	AF275252	AF275252 Xenopus l	c 717	17	1.5	27125	68	AC023654	AC023654 Homo sapi
c 645	17	1.5	2065	94	MUSMHHKA	M11847 Mouse MHC c	c 718	17	1.5	27551	6	CELZC8	U64862 Caenorhabdi
c 646	17	1.5	2072	89	AK025007	AK025007 Homo sapi	c 719	17	1.5	27844	92	HSU17H3	Z84492 Human DNA s
c 647	17	1.5	2103	14	SCREC114X	Z14315 S.cerevisia	c 720	17	1.5	27853	63	AC012639	AC012639 Homo sapi
c 648	17	1.5	2107	6	DOIADCYA03	L05498 Dictyostell	c 721	17	1.5	28992	14	SC9375	Z47071 S.cerevisia
c 649	17	1.5	2123	8	GGU20216	U20216 Gallus gall	c 722	17	1.5	30382	5	CEFA42E11	Z66562 Caenorhabdi
c 650	17	1.5	2210	54	G18283	G18283 SWS1283 Er	c 723	17	1.5	31186	5	CEC32A9	Z49153 Caenorhabdi
c 651	17	1.5	2210	93	HUMACHRM2	M16404 Human m2 mu	c 724	17	1.5	33517	3	SC10B7	AL355752 Streptomy
c 652	17	1.5	2216	85	AB052740S4	AB052743 Homo sapi	c 725	17	1.5	33960	94	HAMBHMC	L12104 Mesocricetu
c 653	17	1.5	2337	94	RATCYP45J	M25143 Rat cytochr	c 726	17	1.5	34182	93	HSU17H3	Z74409 Human DNA s
c 654	17	1.5	2411	8	CHRTKA	K02611 Chicken thy	c 727	17	1.5	34268	60	AC008242	AC008242 Leishmani
c 655	17	1.5	2450	15	SPREP2	X91044 S.pombe rep	c 728	17	1.5	36008	90	AP001230	AP001230 Homo sapi
c 656	17	1.5	2543	92	HSIHGJB1	X86355 H.sapiens i	c 729	17	1.5	36079	5	CEU403E9	AF022981 Caenorhab
c 657	17	1.5	2543	92	HSIHGJB2	X86356 H.sapiens i	c 730	17	1.5	36095	87	AC013256	AC013256 Homo sapi
c 658	17	1.5	2543	92	HSIHGJB3	X86357 H.sapiens i	c 731	17	1.5	36261	5	CELF18A1	U41535 Caenorhabdi
c 659	17	1.5	2543	92	HSIHGJB3	X86358 H.sapiens i	c 732	17	1.5	36471	91	DSU7021	D87021 Homo sapien
c 660	17	1.5	2543	92	HSIHGJC	X86359 H.sapiens i	c 733	17	1.5	36565	93	HSU73648	U73648 Human chrom
c 661	17	1.5	2626	2	AF263927	AF263927 Carnobact	c 734	17	1.5	36649	93	HSU19D8	Z70689 Human DNA s
c 662	17	1.5	2759	6	CEU60113	U60113 Caenorhabdi	c 735	17	1.5	36676	92	HS8588B16	AL096878 Human DNA
c 663	17	1.5	2843	6	CEU60112	U60112 Caenorhabdi	c 736	17	1.5	36776	14	SPAC4G8	Z56276 S.pombe chr
c 664	17	1.5	3096	89	AK026689	AK026689 Homo sapi	c 737	17	1.5	36993	5	CELT10H9	AF067949 Caenorhab
c 665	17	1.5	3182	93	HUMIGCB1	J00256 Human Ig ge	c 738	17	1.5	37319	64	AC014400	AC014400 Drosophil
c 666	17	1.5	3186	88	AF042285	AF042285 Homo sapi	c 739	17	1.5	37388	65	AC017823	AC017823 Drosophil
c 667	17	1.5	3210	89	AK024286	AK024286 Homo sapi	c 740	17	1.5	37773	88	AF027207	AF027207 Homo sapi
c 668	17	1.5	3269	93	HUMQRF1	D26067 Human mRNA	c 741	17	1.5	38012	14	SPC2F12	Z97211 S.pombe chr

742	17	1.5	38345	93	HSUB1830	UB1830 Human cosmi	815	17	1.5	88827	91	HS1J6	AL035652 Human DNA
743	17	1.5	38403	90	AP001241	AP001241 Homo sapi	816	17	1.5	88848	15	T24H34	AF075598 Arabidops
744	17	1.5	38955	6	CEZC84	ZP1957 Caenorhabdi	817	17	1.5	89163	91	HS1009H6	AL035682 Human DNA
745	17	1.5	41134	84	HSUB081	AL022169 Homo sapi	818	17	1.5	89228	64	AC016273	AC016273 Homo sapi
746	17	1.5	41308	85	AC004196	AC004196 Homo sapi	819	17	1.5	89228	64	AC016273	AC016273 Homo sapi
747	17	1.5	41659	60	AC006838	AC006838 Homo sapi	820	17	1.5	90011	90	AL360007	AL360007 Human DNA
748	17	1.5	42569	85	AC005306	AC005306 Homo sapi	821	17	1.5	90770	60	AC011800	AC011800 Homo sapi
749	17	1.5	43218	63	AC012979	AC012979 Drosophi	822	17	1.5	91017	91	HS162C6	Z84475 Human DNA s
750	17	1.5	43952	93	HSU105G4	Z92846 Human DNA s	823	17	1.5	91511	89	AL133404	AL133404 Human DNA
751	17	1.5	47383	9	AX059459	AX059459 Sequence	824	17	1.5	91733	81	AL391278	AL391278 Homo sapi
752	17	1.5	47854	91	HS839M11	AL034485 Human DNA	825	17	1.5	92309	87	AC021908	AC021908 Homo sapi
753	17	1.5	48870	62	AC010838	AC010838 Drosophi	826	17	1.5	92588	92	HSIMMDL	X97051 DNA sequenc
754	17	1.5	50785	12	AC007169	AC007169 Arabidops	827	17	1.5	93338	67	AC022615	AC022615 Homo sapi
755	17	1.5	51041	61	AC008748	AC008748 Homo sapi	828	17	1.5	93626	61	AC010042	AC010042 Drosophi
756	17	1.5	51191	62	AC011870	AC011870 Homo sapi	829	17	1.5	93695	13	ATF2009	AL021749 Arabidops
757	17	1.5	51248	12	AB046437	AB046437 Arabidops	830	17	1.5	93821	67	AC021596	AC021596 Homo sapi
758	17	1.5	51514	66	AC020401	AC020401 Drosophi	831	17	1.5	94817	92	HS981L23	AL031586 Human DNA
759	17	1.5	52691	12	AC007264	AC007264 Arabidops	832	17	1.5	95076	13	ATF7L13	AL049524 Arabidops
760	17	1.5	52706	66	AC019651	AC019651 Drosophi	833	17	1.5	95081	80	AL357138	AL357138 Homo sapi
761	17	1.5	53158	67	AC021375	AC021375 Homo sapi	834	17	1.5	95360	67	AC022465	AC022465 Homo sapi
762	17	1.5	53442	77	AC084432	AC084432 Arabidops	835	17	1.5	95643	14	FN15	AF069299 Arabidops
763	17	1.5	53897	85	AC002037	AC002037 Human Chr	836	17	1.5	95993	13	ATF2A19	AL132962 Arabidops
764	17	1.5	54724	60	AC002555	AC002555 Homo sapi	837	17	1.5	96310	89	AL136444	AL136444 Human DNA
765	17	1.5	55820	81	AL512483	AL512483 Homo sapi	838	17	1.5	96679	13	ATF8P19	AL133315 Arabidops
766	17	1.5	56171	64	AC014872	AC014872 Drosophi	839	17	1.5	96899	12	AC013288	AC013288 Arabidops
767	17	1.5	57620	69	AC024532	AC024532 Homo sapi	840	17	1.5	98188	85	AC005056	AC005056 Homo sapi
768	17	1.5	57871	69	AC025407	AC025407 Homo sapi	841	17	1.5	98844	93	HSWRNG1	AF181896 Homo sapi
769	17	1.5	59828	61	AC009681	AC009681 Homo sapi	842	17	1.5	99814	13	ATF8J2	AL132969 Arabidops
770	17	1.5	61048	78	AC087271	AC087271 Homo sapi	843	17	1.5	99856	13	ATF2G24	AL049488 Arabidops
771	17	1.5	61289	78	AC087536	AC087536 Homo sapi	844	17	1.5	99858	92	HSJ1059A9	AL18524 Human DNA
772	17	1.5	61697	89	AL354933	AL354933 Human DNA	845	17	1.5	100446	78	AF178220	AF178220 Homo sapi
773	17	1.5	62121	71	AC027564	AC027564 Homo sapi	846	17	1.5	100575	4	AC006421	AC006421 Drosophi
774	17	1.5	62268	91	HS1178121	AL109852 Human DNA	847	17	1.5	100806	12	AC015450	AC015450 Arabidops
775	17	1.5	62493	91	HS193G15	AL117391 Human DNA	848	17	1.5	100905	67	AC021758	AC021758 Homo sapi
776	17	1.5	62493	92	HSJ193G15	AL096838 Human DNA	849	17	1.5	101715	13	ATF4D11	AL022537 Arabidops
777	17	1.5	63262	89	AL138968	AL138968 Human DNA	850	17	1.5	102105	89	AL136166	AL136166 Human DNA
778	17	1.5	64837	91	AC027001	AC027001 Homo sapi	851	17	1.5	102552	66	AC019368	AC019368 Drosophi
779	17	1.5	65106	82	AL513533	AL513533 Homo sapi	852	17	1.5	102717	85	AC004079	AC004079 Homo sapi
780	17	1.5	65949	63	AC013340	AC013340 Homo sapi	853	17	1.5	103282	81	AL445219	AL445219 Homo sapi
781	17	1.5	67916	78	AC087280	AC087280 Homo sapi	854	17	1.5	103563	86	AC006979	AC006979 Homo sapi
782	17	1.5	68753	87	AC008822	AC008822 Homo sapi	855	17	1.5	104835	88	AF130418	AF130418 Homo sapi
783	17	1.5	68790	92	HSJ287H17	AL121970 Human DNA	856	17	1.5	105350	82	AF000661	AF000661 Homo sapi
784	17	1.5	69060	85	AC005196	AC005196 Homo sapi	857	17	1.5	105357	91	HS34P24	Z82202 Human DNA s
785	17	1.5	69550	12	AC002388	AC002388 Arabidops	858	17	1.5	105357	91	HS34P24	Z82202 Human DNA s
786	17	1.5	71734	85	AC004756	AC004756 Homo sapi	859	17	1.5	106230	91	HS629B2	AL035412 Human DNA
787	17	1.5	72758	74	AC069108	AC069108 Homo sapi	860	17	1.5	106935	6	CELY37E3	AC087079 Caenorhab
788	17	1.5	73844	92	HSBK109D1	AC079300 Human DNA	861	17	1.5	107469	91	HS29H4	AL078475 Homo sapi
789	17	1.5	73977	13	AP000417	AP000417 Arabidops	862	17	1.5	107839	82	AF000574	AF000574 Homo sapi
790	17	1.5	74531	64	AC015829	AC015829 Homo sapi	863	17	1.5	107935	64	AC016624	AC016624 Homo sapi
791	17	1.5	74747	4	AC084466	AC084466 Caenorhab	864	17	1.5	108040	73	AC068150	AC068150 Homo sapi
792	17	1.5	75961	67	AC022653	AC022653 Homo sapi	865	17	1.5	109018	92	HSJ1129A6	AL096763 Human DNA
793	17	1.5	76072	14	NCB1D1	AL355927 Neurospor	866	17	1.5	110000	60	AC008653_2	Continuation (3 of
794	17	1.5	79419	92	HSJ324O17	AL110115 Human DNA	867	17	1.5	110000	61	AC009579_1	Continuation (2 of
795	17	1.5	79528	91	HS466N1	Z97630 Human DNA s	868	17	1.5	110000	80	AL355364_3	Continuation (4 of
796	17	1.5	79914	85	AC005159	AC005159 Homo sapi	869	17	1.5	110000	80	AL356097_2	Continuation (3 of
797	17	1.5	80279	62	AC011624	AC011624 Arabidops	870	17	1.5	110000	80	AL359267_2	Continuation (3 of
798	17	1.5	81001	9	AX063464	AX063464 Sequence	871	17	1.5	110000	83	CEY39A1_0	Z99293 Caenorhabdi
799	17	1.5	81171	64	AC016218	AC016218 Homo sapi	872	17	1.5	110000	83	CEY39A1_1	Continuation (2 of
800	17	1.5	81171	64	AC016218	AC016218 Homo sapi	873	17	1.5	110079	94	AC002315	AC002315 Mouse BAC
801	17	1.5	81295	12	AC006161	AC006161 Arabidops	874	17	1.5	110085	87	AC016622	AC016622 Homo sapi
802	17	1.5	82001	12	AB005242	AB005242 Arabidops	875	17	1.5	110870	91	HS154P24	Z94044 Human DNA s
803	17	1.5	82053	12	AC022455	AC022455 Arabidops	876	17	1.5	110960	6	CEY39A1A	AL031633 Caenorhab
804	17	1.5	82160	68	AC023027	AC023027 Homo sapi	877	17	1.5	111126	88	AF321237	AF321237 Homo sapi
805	17	1.5	82279	90	AL391071	AL391071 Human DNA	878	17	1.5	112519	85	AB041992	AB041992 Homo sapi
806	17	1.5	82517	91	HS341D10	Z97985 Human DNA s	879	17	1.5	113060	60	AC008255	AC008255 Drosophi
807	17	1.5	83057	90	AL445703	AL445703 Human DNA	880	17	1.5	113761	61	AC008796	AC008796 Homo sapi
808	17	1.5	83078	12	AF072897	AF072897 Arabidops	881	17	1.5	113980	82	AP001540	AP001540 Homo sapi
809	17	1.5	83569	68	AC022967	AC022967 Homo sapi	882	17	1.5	114478	82	AF000489	AF000489 Homo sapi
810	17	1.5	83859	12	AB024038	AB024038 Arabidops	883	17	1.5	116561	91	HS753P9	AL023653 Human DNA
811	17	1.5	85567	12	AB016878	AB016878 Arabidops	884	17	1.5	116585	78	AL139153	AL139153 Homo sapi
812	17	1.5	87180	12	AB023044	AB023044 Arabidops	885	17	1.5	116692	92	HSJ570012	AL121835 Human DNA
813	17	1.5	87219	13	AP002054	AP002054 Arabidops	886	17	1.5	117283	83	CNS01RIL	AL163639 Homo sapi
814	17	1.5	87937	83	CNS05TEA	AL358339 Homo sapi	887	17	1.5	117296	12	AC008153	AC008153 Arabidops

c 888	17	1.5	117824	90	AL389886	Human DNA	AL389886	Human DNA	961	17	1.5	140270	63	AC012602	Homo sapi	
c 889	17	1.5	118061	75	AC073598	Homo sapi	AC073598	Homo sapi	c 962	17	1.5	140426	81	AL365216	Homo sapi	
c 890	17	1.5	118212	60	AC007581	Drosophil	AC007581	Drosophil	c 963	17	1.5	140497	61	AC008814	Homo sapi	
c 891	17	1.5	118695	77	AC084056	Mus muscu	AC084056	Mus muscu	c 964	17	1.5	140832	60	AC008212	Drosophil	
c 892	17	1.5	119119	87	AC008970	Homo sapi	AC008970	Homo sapi	c 965	17	1.5	141387	86	AC005686	Homo sapi	
c 893	17	1.5	119182	92	HSB62120	Homo sapi	AL050302	Homo sapi	c 966	17	1.5	141642	70	AC025836	Homo sapi	
c 894	17	1.5	120188	64	AC014616	Homo sapi	AC014616	Homo sapi	c 967	17	1.5	141824	71	AC027675	Homo sapi	
c 895	17	1.5	120507	89	AL136358	Human DNA	AL136358	Human DNA	c 968	17	1.5	141894	87	AC008833	Homo sapi	
c 896	17	1.5	121430	60	AC008355	Drosophil	AC008355	Drosophil	c 969	17	1.5	142166	60	AC008405	Homo sapi	
c 897	17	1.5	121501	12	AC069557	Genomic S	AC069557	Genomic S	c 970	17	1.5	142190	67	AC021415	Homo sapi	
c 898	17	1.5	122000	85	AC003093	Human BAC	AC003093	Human BAC	c 971	17	1.5	142650	66	AC019040	Homo sapi	
c 899	17	1.5	122764	92	HSJ831D17	Human DNA	AL109984	Human DNA	c 972	17	1.5	142665	91	HS202121	Human DNA	
c 900	17	1.5	122979	92	HSBK14H9	Human DNA	AL121936	Human DNA	c 973	17	1.5	142723	64	AC016517	Homo sapi	
c 901	17	1.5	122986	85	AC004915	Homo sapi	AC004915	Homo sapi	c 974	17	1.5	142942	62	AC010501	Homo sapi	
c 902	17	1.5	123209	89	AL158195	Human DNA	AL158195	Human DNA	c 975	17	1.5	143027	80	AL360086	Homo sapi	
c 903	17	1.5	123300	82	AP000406	Homo sapi	AP000406	Homo sapi	c 976	17	1.5	143068	93	HSU95626	Homo sapi	
c 904	17	1.5	123593	80	AL356912	Homo sapi	AL356912	Homo sapi	c 977	17	1.5	143107	86	AC005949	Homo sapi	
c 905	17	1.5	123846	92	HSJ596C3	Human DNA	AL096827	Human DNA	c 978	17	1.5	143561	72	AC046191	Homo sapi	
c 906	17	1.5	124660	82	AP001954	Homo sapi	AP001954	Homo sapi	c 979	17	1.5	143879	12	AC069251	Genomic S	
c 907	17	1.5	124764	60	AC008377	Homo sapi	AC008377	Homo sapi	c 980	17	1.5	144015	80	AL357144	Homo sapi	
c 908	17	1.5	125727	60	AC007928	Drosophil	AC007928	Drosophil	c 981	17	1.5	144214	60	AC008211	Drosophil	
c 909	17	1.5	125933	82	AP000709	Homo sapi	AP000709	Homo sapi	c 982	17	1.5	144452	80	AL357895	Homo sapi	
c 910	17	1.5	126546	90	AP000689	Homo sapi	AP000689	Homo sapi	c 983	17	1.5	144514	71	AC027433	Homo sapi	
c 911	17	1.5	126790	69	AC025180	Homo sapi	AC025180	Homo sapi	c 984	17	1.5	144585	70	AC025736	Homo sapi	
c 912	17	1.5	127277	87	AC010175	Homo sapi	AC010175	Homo sapi	c 985	17	1.5	144623	78	AL136322	Homo sapi	
c 913	17	1.5	128142	89	AL356479	Human DNA	AL356479	Human DNA	c 986	17	1.5	145626	66	AC020721	Homo sapi	
c 914	17	1.5	128206	82	AP001258	Homo sapi	AP001258	Homo sapi	c 987	17	1.5	145734	62	AC011096	Homo sapi	
c 915	17	1.5	129813	61	AC008823	Homo sapi	AC008823	Homo sapi	c 988	17	1.5	146017	71	AC027473	Homo sapi	
c 916	17	1.5	129914	85	AC003070	Homo sapi	AC003070	Homo sapi	c 989	17	1.5	146098	88	AC074112	Homo sapi	
c 917	17	1.5	130757	76	AC079584	Homo sapi	AC079584	Homo sapi	c 990	17	1.5	146106	79	AL353894	Homo sapi	
c 918	17	1.5	131329	79	AL354887	Homo sapi	AL354887	Homo sapi	c 991	17	1.5	146244	82	AP002083	Homo sapi	
c 919	17	1.5	131394	80	AL356600	Homo sapi	AL356600	Homo sapi	c 992	17	1.5	146258	68	AC023479	Homo sapi	
c 920	17	1.5	131684	62	AC010577	Drosophil	AC010577	Drosophil	c 993	17	1.5	146333	86	AC007680	Homo sapi	
c 921	17	1.5	131735	91	HS50A13	Human DNA	Z92545	Human DNA	c 994	17	1.5	146335	63	AC012626	Homo sapi	
c 922	17	1.5	132010	69	AC024351	Homo sapi	AC024351	Homo sapi	c 995	17	1.5	146574	68	AC023459	Homo sapi	
c 923	17	1.5	132055	79	AL161742	Homo sapi	AL161742	Homo sapi	c 996	17	1.5	146596	90	AP000497	Homo sapi	
c 924	17	1.5	132150	86	AC005586	Homo sapi	AC005586	Homo sapi	c 997	17	1.5	146596	90	AP000497	Homo sapi	
c 925	17	1.5	132727	80	AL356914	Homo sapi	AL356914	Homo sapi	c 998	17	1.5	146644	66	AC019175	Homo sapi	
c 926	17	1.5	132910	64	AC016342	Homo sapi	AC016342	Homo sapi	c 999	17	1.5	147288	78	AC083861	Homo sapi	
c 927	17	1.5	133508	88	AF241728	Homo sapi	AF241728	Homo sapi	c 1000	17	1.5	147432	80	AC026467	Homo sapi	
c 928	17	1.5	133574	91	HS20N2	Human DNA	AL031320	Human DNA				1.5	147495	80	AL358781	Homo sapi
c 929	17	1.5	133747	89	AL139150	Human DNA	AL139150	Human DNA								
c 930	17	1.5	133787	87	AC015540	Homo sapi	AC015540	Homo sapi								
c 931	17	1.5	134040	72	AC044856	Homo sapi	AC044856	Homo sapi								
c 932	17	1.5	134142	80	AL359976	Homo sapi	AL359976	Homo sapi								
c 933	17	1.5	134222	85	AB045363	Homo sapi	AB045363	Homo sapi								
c 934	17	1.5	134333	84	HSJ636L22	Homo sapi	AL109844	Homo sapi								
c 935	17	1.5	134609	67	AC021416	Homo sapi	AC021416	Homo sapi								
c 936	17	1.5	134884	64	AC016600	Homo sapi	AC016600	Homo sapi								
c 937	17	1.5	134952	78	AL139104	Homo sapi	AL139104	Homo sapi								
c 938	17	1.5	135259	85	AC004590	Homo sapi	AC004590	Homo sapi								
c 939	17	1.5	135569	75	AC073997	Homo sapi	AC073997	Homo sapi								
c 940	17	1.5	135736	12	AC037426	Oryza sat	AC037426	Oryza sat								
c 941	17	1.5	135928	91	HS434014	Homo sapi	AL022398	Homo sapi								
c 942	17	1.5	136131	85	AC000360	Homo sapi	AC000360	Homo sapi								
c 943	17	1.5	136333	67	AC022015	Homo sapi	AC022015	Homo sapi								
c 944	17	1.5	136551	72	AC048354	Human DNA	AC048354	Human DNA								
c 945	17	1.5	136649	92	HSJ635E8	Human DNA	AL110502	Human DNA								
c 946	17	1.5	136725	88	AC026369	Homo sapi	AC026369	Homo sapi								
c 947	17	1.5	136742	2	ECU82598	Escherichia	U82598	Escherichia								
c 948	17	1.5	136750	80	AL359884	Homo sapi	AL359884	Homo sapi								
c 949	17	1.5	137216	81	AL450244	Homo sapi	AL450244	Homo sapi								
c 950	17	1.5	137347	89	AL354674	Human DNA	AL354674	Human DNA								
c 951	17	1.5	137481	71	AC027207	Homo sapi	AC027207	Homo sapi								
c 952	17	1.5	137833	83	CNS01DV6	Homo sapi	AL133455	Homo sapi								
c 953	17	1.5	137940	64	AC015612	Homo sapi	AC015612	Homo sapi								
c 954	17	1.5	138443	91	HS179E13	Human DNA	AL035603	Human DNA								
c 955	17	1.5	138868	67	AC022375	Homo sapi	AC022375	Homo sapi								
c 956	17	1.5	139049	85	AC004147	Homo sapi	AC004147	Homo sapi								
c 957	17	1.5	139113	70	AC026372	Homo sapi	AC026372	Homo sapi								
c 958	17	1.5	139218	66	AC019289	Homo sapi	AC019289	Homo sapi								
c 959	17	1.5	140056	60	AC007863	Trypanoso	AC007863	Trypanoso								
c 960	17	1.5	140186	73	AC055808	Homo sapi	AC055808	Homo sapi								

## ALIGNMENTS

RESULT	1	HSCKRL3	1255 bp	DNA	PRI
LOCUS		HSCKRL3			13-NOV-1996
DEFINITION		H.sapiens G protein-coupled receptor CKR-L3.			
ACCESSION		Z79784			
VERSION		Z79784.1			
KEYWORDS		G Protein-coupled Receptor CKR-L3.			
SOURCE		human.			
ORGANISM		Homo sapiens			
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;			
AUTHORS		1 (bases 1 to 1255)			
TITLE		Gutierrez, J., Varona, R., Zaballos, A., Lind, P. and Marquez, G.			
JOURNAL		unpublished			
REFERENCE		2 (bases 1 to 1255)			
AUTHORS		Zaballos, A.			
TITLE		Direct Submission			
JOURNAL		Submitted (03-SEP-1996) Angel Zaballos, Research, Pharmacia &			
REFERENCE		3 (bases 1 to 1255)			
AUTHORS		Zaballos, A., Varona, R., Gutierrez, J., Lind, P. and Marquez, G.			
TITLE		Molecular cloning and RNA expression of two new human chemokine			
JOURNAL		receptor-like genes			
MEDLINE		Biochem. Biophys. Res. Commun. 227 (3), 846-853 (1996)			
REMARK		Erratum: [[published erratum appears in Biochem Biophys Res Commun			
		1997 Feb 13;231(2):519-20]]			

FEATURES	Location/Qualifiers
source	1. .1255
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="6"
CDS	76..1185
	/codon_start=1
	/product="G PROTEIN-COUPLED RECEPTOR CKR-L3"
	/protein_id="CAB02144.1"
	/db_xref="GI:1668738"
	/db_xref="SWISS-PROT:P51684"
	/translation="MNFSVDFSDSEDFVSVNTSYSDSEMLLCSLQVRFSLFV PLAYSILCVGLGLNLLVYITFAFKKARSMTDVLLNMAIADILFVLPLFWAVSHA TGAVFNSATKLLIGIYAINFCNGMLLTICISMDRYTAIVQATSKFRLSRILPRSK IICLVWGLSVIISSSTFVNKYNTQSDCEPKYQTVSEPIRWKLLMLGLLELFGF FLPLMFIPCYFTFIVKTLVOAONSKRKAIRVIAVLVFLACQIPHNMLLVTAANL GKMNRSCEKELIGYKTVLEAFLLHCLCNPLVAFIQKERNFLKILKDLWCVR KYSKGFSCAGRYSENISQTSADNDNASFTM"
BASE COUNT	302 a 294 c 284 g 375 t
ORIGIN	
Query Match 77.7%; Score 869; DB 92; Length 1255;	
Best Local Similarity 99.8%; Pred. No. 0;	
Matches 969; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Qy	149 TTGGCCCTCTGGGGAATATCTGGTGTGATCACCTTGTCTTTTATAGAAGCCAGGT 208
Db	
Qy	209 CTATGACAGAGCTATCTCTTGAACATGGCCATTCGACACATCCCTCTTGTCTTACTC 268
Db	
Qy	269 TCCCATCTCGGAGAGTATGACACAGTGGTGGGTTTTCACCAATGCCAGTGCA 328
Db	
Qy	329 AGTTGCTAAAAGGCATCTATGCGCATCAACTTAACTGCGGAGTGTCTCTCACTTGCA 388
Db	
Qy	389 TTAGCATGGACCGGTATCGCATTTGACAGCGACTAAGTCATTCGCGCTCCGATCCA 448
Db	
Qy	449 GAACACTACCGCAGCAAAATCATCTGCTTGTGTGGGGCTGTGAGTCATCATCT 508
Db	
Qy	509 CCAGCTCAACTTTGTCTTCAACCAAAAATACAAACCCCAAGCAGCGATGCTGTGAAC 568
Db	
Qy	569 CCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGAAGTGTGTGTTGGGGCTTGAGC 628
Db	
Qy	629 TACTCTTTGGTTCTTTATCCCTTTGATGTTCAATATATTTTGTACAGTTCAATGTCA 688
Db	
Qy	716 TACTCTTTGGTTCTTTATCCCTTTGATGTTCAATATATTTTGTACAGTTCAATGTCA 775
Db	
Qy	749 TGGTGTGTGTCTGTGGTGTGTCAGATTCCTATACATGTCCTGTGTGACGGCTG 808
Db	
Qy	836 TGGTGTGTGTCTGTGGTGTGTCAGATTCCTATACATGTCCTGTGTGACGGCTG 895
Db	
Qy	809 CTAATTTGGTAAATGAACCCGATCCTGCCAGAGCGAAAGCTAAATGGCTATACGAAA 868
Db	
Qy	896 CAAATTTGGTAAATGAACCCGATCCTGCCAGAGCGAAAGCTAAATGGCTATACGAAA 955
Db	

QY	869	CTGTCACAGAAGTCTCTGGCTTTCTCTGCACGTGCTGCCGTGAACCCCTGCTGTCTACGCTTTTA	928
Db	956	CTGTCACAGAAGTCTCTGGCTTTCTCTGCACGTGCTGCCGTGAACCCCTGCTGTCTACGCTTTTA	1015
QY	929	TTGGGCGAGAAGTCTCAGAACTACTTCTTGAGATCTTTGAAGACCTCTGCTGTGTGAGAA	988
Db	1016	TTGGGCGAGAAGTCTCAGAACTACTTCTTGAGATCTTTGAAGACCTCTGCTGTGTGAGAA	1075
QY	989	GGAAGTACAACTCTCAGGCTTCTCCTGTGCGGGAGGTACTCAGAAAACATTTCTCGGC	1048
Db	1076	GGAAGTACAACTCTCAGGCTTCTCCTGTGCGGGAGGTACTCAGAAAACATTTCTCGGC	1135
QY	1049	AGACCACTGAGACCGCAGATAACGACAAATCGCTGCTCTTCTACTATGTATAGAAAAGCTG	1108
Db	1136	AGACCACTGAGACCGCAGATAACGACAAATCGCTGCTCTTCTACTATGTATAGAAAAGCTG	1195
QY	1109	AGTCTCCCTTAA 1119	
Db	1196	AGTCTCCCTTAA 1206	
RESULT	2		
LOCUS	HSU45984		
DEFINITION	Homo sapiens CCR6 chemokine receptor (CMKBR6) gene, complete cds.		
ACCESSION	U45984		
VERSION	U45984.1		
KEYWORDS	GI:2246432		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 3693)		
JOURNAL	Baba, H., Imai, T., Nishimura, M., Kakizaki, M., Takagi, S.,		
MEDLINE	Hieshima, K., Nomiya, H. and Yoshie, O.		
REFERENCE	Identification of CCR6, the specific receptor for a novel		
AUTHORS	lymphocyte-directed CC chemokine LARC		
TITLE	J. Biol. Chem. 272 (23), 14893-14898 (1997)		
JOURNAL	97313465		
MEDLINE	2 (bases 1 to 3693)		
REFERENCE	Lautens, L. L., Modi, W. and Bonner, T. I.		
AUTHORS	Cloning, Tissue Distribution and Chromosomal Localization of a		
TITLE	potential G-Protein-Linked Receptor		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 3693)		
AUTHORS	Bonner, T. I.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JAN-1996) Tom I. Bonner, Lab of Cell Biology, NIMH,		
COMMENT	Bldg 36, Rm 3A-17, MSC 4090, Bethesda, MD 20892-4090, USA		
FEATURES	On Jul 8, 1997 this sequence version replaced gi:1245058.		
source	Location/Qualifiers		
	1. .3693		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="6"		
	/map="6q27"		
	/clone="GPR-CY4"		
	join(35..140,237..3136)		
	132..1352		
	/gene="CMKBR6"		
	join(132..140,237..1352)		
	/gene="CMKBR6"		
	/note="G protein-coupled receptor"		
	/codon_start=1		
	/product="CCR6 chemokine receptor"		
	/protein_id="AAB62714.1"		
	/db_xref="GI:2251211"		
	/translation="MSGESMNSDFSDSEDFVSVNTSYSDSEMLLCSLQVRFSLFV SRILFPIAYSLFCVGLGLNLLVYITFAFKKARSMTDVLLNMAIADILFVLPLFWAVSHA LYSHATGAWFNSATKLLIGIYAINFCNGMLLTICISMDRYTAIVQATSKFRLSRILPRSK IICLVWGLSVIISSSTFVNKYNTQSDCEPKYQTVSEPIRWKLLMLGLLELFGF FLPLMFIPCYFTFIVKTLVOAONSKRKAIRVIAVLVFLACQIPHNMLLVTAANL GKMNRSCEKELIGYKTVLEAFLLHCLCNPLVAFIQKERNFLKILKDLWCVR KYSKGFSCAGRYSENISQTSADNDNASFTM"		
mRNA			
gene			
CDS			









```
QY 689 AAACCTTGGTCAAGCTCAGAATTCATAAAGCCACAAAGCCATCGTGTAATCATATAGCTG 748
Db 980 AAACCTTGGTCAAGCTCAGAATTCATAAAGCCACAAAGCCATCGTGTAATCATATAGCTG 1039
QY 749 TGGTGTCTTGTGTTCTGGCTTGTCTAGATTCCTCATAAATGCTCCTGCTGTGTGACGGCTG 808
Db 1040 TGGTGTCTTGTGTTCTGGCTTGTCTAGATTCCTCATAAATGCTCCTGCTGTGTGACGGCTG 1099
QY 809 CTAATTTGGGTAATAATGAACCGATCCTGCCAGAGCGAAAGCTAATTTGGCTATACGAAA 868
Db 1100 CAAATTTGGGTAATAATGAACCGATCCTGCCAGAGCGAAAGCTAATTTGGCTATACGAAA 1159
QY 869 CTGTACAGAACTCTGGCTTCTCTGCACTGTGCTGCAACCTGTGCTCTAGCTTTTA 928
Db 1160 CTGTACAGAACTCTGGCTTCTCTGCACTGTGCTGCAACCTGTGCTCTAGCTTTTA 1219
QY 929 TTGGGCAGAAAGTTCAAGAACTACTTTCTGAAGATCTTGAAGACCTGTGCTGTGAGAA 988
Db 1220 TTGGGCAGAAAGTTCAAGAACTACTTTCTGAAGATCTTGAAGACCTGTGCTGTGAGAA 1279
QY 989 GGAAGTACAAGTCTTCAGGCTTCTCTGTCGGGAGGTACTCAGAAAACATTTCTCGGC 1048
Db 1280 GGAAGTACAAGTCTTCAGGCTTCTCTGTCGGGAGGTACTCAGAAAACATTTCTCGGC 1339
QY 1049 AGACCAAGTACAGCCAGATACAGCAATAGCTGCTGCTCTCTACTATGTGATAGAAAAGCTG 1108
Db 1340 AGACCAAGTACAGCCAGATACAGCAATAGCTGCTGCTCTCTACTATGTGATAGAAAAGCTG 1399
QY 1109 AGTCTCCCTAA 1119
Db 1400 AGTCTCCCTAA 1410

RESULT 6
LOCUS H5U60000
DEFINITION Human IL8-related receptor (DRY6) mRNA, complete cds.
ACCESSION U60000
VERSION U60000.1 GI:1515434
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1137)
AUTHORS McCoy,R. and Perlmutter,D.H.
TITLE Cloning of novel IL8-related receptors from hepatic tissue
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1137)
AUTHORS McCoy,R. and Perlmutter,D.H.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1996) Pediatrics, Washington University in St.
Louis, 4942 Parkview Place, St Louis, MO 63110, USA

FEATURES
source
1..1137
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="hepatoma HepG2"
1..1137
/gene="DRY6"
1..1137
/gene="DRY6"
/codon_start=1
/product="IL8-related receptor"
/protein_id="AAB06949.1"
/db_xref="GI:1515435"
/translation="MSGESMNFSDVFDSSDDYFVSVNTSYSDSEMLLCSLQEVRFQF
SRLEVPVAYSLICVFALLGNILVITFAFNKARKSMTDVIYLVNMAIADILFVLTLPFW
AVSHATGAWFNSATCKLLGIYAINFCGMKLLTICISMDRYIAIVQATKSPRLRSRT
LPRSKICLVVWGLSVIISSSSFVNOKYNTLGSDDVCEPKYQTVSPVIRWKLMLGLE
LLGFFFPFLMEIFCYFIVTKTLVQAKNSKHKAIRVILIAVLVFLACQIPHNWVLL
TAANLGRNKRSCSEKLLIGYTKTVTEVLAFDHCLLPVLYAFIQGKFRNRYFLKLDKL
```

```
BASE COUNT 263 a 270 c 266 g 338 t
ORIGIN
Query Match 39.5%; Score 442; DB 93; Length 1137;
Best Local Similarity 99.1%; Pred. No. 7.7e-226;
Matches 842; Conservative 0; Mismatches 8; Indels 0; Gaps 0:
QY 230 TGAACATGGCCATTGCAGACATCCTCTTTGTCTTACTCTCCCATTTCTGGCAGTGAAGTC 289
Db 257 TGAACATGGCCATTGCAGACATCCTCTTTGTCTTACTCTCCCATTTCTGGCAGTGAAGTC 316
QY 290 ATGCCACTGTGCTGGGTTTTTCAGCAATGCCAGTGCAGAACTTGTCTAAAAGCATCTATG 349
Db 317 ATGCCACTGTGCTGGGTTTTTCAGCAATGCCAGTGCAGAACTTGTCTAAAAGCATCTATG 376
QY 350 CCATCAACTTTAACTGCGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 409
Db 377 CCATCAACTTTAACTGCGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
QY 410 CCATTTGACAGGCGACTAAGTCAATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAA 469
Db 437 CCATTTGACAGGCGACTAAGTCAATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAA 496
QY 470 TCATCTGCCCTTGTGTGTGGGGCTGTGAGTCATCATCTCCAGTCAACTTTTGTCTTCA 529
Db 497 TCATCTGCCCTTGTGTGTGGGGCTGTGAGTCATCATCTCCAGTCAACTTTTGTCTTCA 556
QY 530 ACCAAAATAACACACCAAGGCGAGCGATGCTGTGACCCCAAGTACCAAACTGCTCTGG 589
Db 557 ACCAAAATAACACACCAAGGCGAGCGATGCTGTGACCCCAAGTACCAAACTGCTCTGG 616
QY 590 AGCCCATCAGGTGGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGGTTCTTTATCC 649
Db 617 TGCCCATCAGGTGGAAGCTGCTGATGTTGGGGCTTGAGCTACTCTTTGGTTCTTTATCC 676
QY 650 CTTTGATGTTCAATGATATTTTGTACAGTTCATGTCAAAACCTTGGTGCAGCTCAGA 709
Db 677 CTTTGATGTTCAATGATATTTTGTACAGTTCATGTCAAAACCTTGGTGCAGCTCAGA 736
QY 710 ATTCTAAAGGCACAAAGCCATCGTGTATCATAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 769
Db 737 ATTCTAAAGGCACAAAGCCATCGTGTATCATAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 796
QY 770 GTGAGATTCCTCATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 829
Db 797 GTGAGATTCCTCATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
QY 830 GATCTGCCAGAGCGGAAAGCTAATTTGGCTATACGAAACTGTACAGAACTGCTGCTGCTGCTG 889
Db 857 GATCTGCCAGAGCGGAAAGCTAATTTGGCTATACGAAACTGTACAGAACTGCTGCTGCTGCTG 916
QY 890 TCCTGCACCTGCTGCTCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 949
Db 917 TCCTGCACCTGCTGCTCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
QY 950 ACTTTCTGAAGATCTTTGAAGGACCTGTGGTGTGTGAGAAGGAGTACAAAGTCTCAGGCT 1009
Db 977 ACTTTCTGAAGATCTTTGAAGGACCTGTGGTGTGTGAGAAGGAGTACAAAGTCTCAGGCT 1036
QY 1010 TCTCCTGTGCCGGAGGTACTCAGAAAACATTTCTCGGAGACCAGTGCAGACCCAGATA 1069
Db 1037 TCTCCTGTGCCGGAGGTACTCAGAAAACATTTCTCGGAGACCAGTGCAGACCCAGATA 1096
QY 1070 ACCACAATGC 1079
Db 1097 ACCACAATGC 1106

RESULT 7
LOCUS AL353591/c 205272 bp DNA HTG 20-JAN-2001
```

DEFINITION Homo sapiens chromosome 6 clone RP11-366H19, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 18 unordered pieces.  
ACCESSION AL353591  
VERSION AL353591.5 GI:9930866  
KEYWORDS HTG: HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 205272)  
Sims,S.  
Direct Submission  
Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Aug 27, 2000 this sequence version replaced gi:9863646.  
----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: BA366H19  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: plasmid; L08752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 192563 bases at least Q40  
Consensus quality: 197287 bases at least Q30  
Consensus quality: 200442 bases at least Q20  
Insert size: 203572; sum-of-contigs  
Insert size: 183107; 13.0% error; agarose-fp  
Quality coverage: 3.30x in Q20 bases; sum-of-contigs Quality  
coverage: 3.75x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 18 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 4259: contig of 4259 bp in length  
\* 4260 4359: gap of 100 bp  
\* 4360 26102: contig of 21743 bp in length  
\* 26103 26202: gap of 100 bp  
\* 26203 28547: contig of 2345 bp in length  
\* 28548 28647: gap of 100 bp  
\* 28648 49653: contig of 21006 bp in length  
\* 49654 49753: gap of 100 bp  
\* 49754 80134: contig of 30381 bp in length  
\* 80135 80234: gap of 100 bp  
\* 80235 83646: contig of 3412 bp in length  
\* 83647 83746: gap of 100 bp  
\* 83747 96061: contig of 12315 bp in length  
\* 96062 96161: gap of 100 bp  
\* 96162 103311: contig of 7150 bp in length  
\* 103312 103411: gap of 100 bp  
\* 103412 122508: contig of 19097 bp in length  
\* 122509 122608: gap of 100 bp  
\* 122609 124966: contig of 2358 bp in length  
\* 124967 125066: gap of 100 bp  
\* 125067 142955: contig of 17889 bp in length  
\* 142956 143055: gap of 100 bp  
\* 143056 161736: contig of 18681 bp in length  
\* 161737 161836: gap of 100 bp  
\* 161837 169189: contig of 7353 bp in length  
\* 169190 169289: gap of 100 bp  
\* 169290 175251: contig of 5962 bp in length  
\* 175252 175351: gap of 100 bp  
\* 175352 181328: contig of 5977 bp in length  
\* 181329 181428: gap of 100 bp

\* 181429 183863: contig of 2435 bp in length  
\* 183864 183963: gap of 100 bp  
\* 183964 195169: contig of 11206 bp in length  
\* 195170 195269: gap of 100 bp  
\* 195270 205272: contig of 10003 bp in length.  
FEATURES  
Location/Qualifiers  
source  
1..205272  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="RP11-366H19"  
/clone\_lib="RPC1-11.2"  
1..4259  
/note="assembly\_fragment:00710  
fragment\_chain:1"  
4360..26102  
/note="assembly\_fragment:00448  
fragment\_chain:1"  
26203..28547  
/note="assembly\_fragment:01871  
fragment\_chain:1"  
28648..49653  
/note="assembly\_fragment:01605  
fragment\_chain:1"  
49754..80134  
/note="assembly\_fragment:01641  
fragment\_chain:1"  
80235..83646  
/note="assembly\_fragment:01500  
fragment\_chain:1"  
83747..96061  
/note="assembly\_fragment:01099  
fragment\_chain:1"  
96162..103311  
/note="assembly\_fragment:00756  
fragment\_chain:2"  
103412..122508  
/note="assembly\_fragment:01078  
fragment\_chain:2"  
122609..124966  
/note="assembly\_fragment:01675  
fragment\_chain:2"  
125067..142955  
/note="assembly\_fragment:01165  
fragment\_chain:3"  
143056..161736  
/note="assembly\_fragment:01021  
fragment\_chain:3"  
161837..169189  
/note="assembly\_fragment:01200  
fragment\_chain:3"  
169290..175251  
/note="assembly\_fragment:00119"  
175352..181328  
/note="assembly\_fragment:00401"  
181429..183863  
/note="assembly\_fragment:01812"  
183964..195169  
/note="assembly\_fragment:02127"  
195270..205272  
/note="assembly\_fragment:01692  
clone\_end:T7  
vector\_side:right"  
BASE COUNT 53785 a 47341 c 47491 g 54940 t 1715 others  
ORIGIN  
Query Match 38.8%; Score 434; DB 79; Length 205272;  
Best Local Similarity 99.6%; Pred. No. 5.6e-222;  
Matches 534; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 149 TTGGCCTCTGGGGAATATTCGTGGTGATCACCTTTGCTTTTATAAGAGCCAGGT 208  
|||||

```
Db 123144 TTGGCTCTCTGGGGAATATCTCGTGGTGATCATCACTTTGCTTTTATAGAAGCCAGGT 123085
QY 209 CTATGACAGACGCTATCTCTTGAACATGCGCATTCGAGACATCCCTCTTTGTTCTTACTC 268
|||||
Db 123084 CTATGACAGACGCTATCTCTTGAACATGCGCATTCGAGACATCCCTCTTTGTTCTTACTC 123025
QY 269 TCCCATCTTGGGCGAGTGAGTCATGCCACTGGTCGGTGGGTTTTCAGCAATGCCACGTGCA 328
|||||
Db 123024 TCCCATCTTGGGCGAGTGAGTCATGCCACTGGTCGGTGGGTTTTCAGCAATGCCACGTGCA 122965
QY 329 AGTTGCTAAAGGCATCTATGCGCATCAACTTTAACTGCGGGATGCTGCTCTGACTGTGCA 388
|||||
Db 122964 AGTTGCTAAAGGCATCTATGCGCATCAACTTTAACTGCGGGATGCTGCTCTGACTGTGCA 122905
QY 389 TTAGCATGACCGGTATCATCGCCATTGTACAGCGCACTAAAGTCATTCGCGCTCCGATCCA 448
|||||
Db 122904 TTAGCATGACCGGTATCATCGCCATTGTACAGCGCACTAAAGTCATTCGCGCTCCGATCCA 122845
QY 449 GAACACTACCGCGCACAAATCATCTGCCCTTGTGTGGGGCTGTCAAGTCATCATCT 508
|||||
Db 122844 GAACACTACCGCGCACAAATCATCTGCCCTTGTGTGGGGCTGTCAAGTCATCATCT 122785
QY 509 CCAGCTCAACTTTTGTCTTCAACCAAAATACACACCCCAAGCAGCATGCTGTGTAAC 568
|||||
Db 122784 CCAGCTCAACTTTTGTCTTCAACCAAAATACACACCCCAAGCAGCATGCTGTGTAAC 122725
QY 569 CCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGAAGCTGCTGATGTTGGGCTTGAGC 628
|||||
Db 122724 CCAAGTACCAAACTGTCTCGGAGCCCATCAGGTGGAAGCTGCTGATGTTGGGCTTGAGC 122665
QY 629 TACTCTTTGGTTCTTATCCCTTTGATGTTTCATGATATTTTTCACACCTTCATT 684
|||||
Db 122664 TACTCTTTGGTTCTTATCCCTTTGATGTTTCATGATATTTTTCACACCTTCATT 122609

RESULT 8
AL390375/c
LOCUS
DEFINITION Homo sapiens chromosome 13 clone RP11-505I19, *** SEQUENCING IN
PROGRESS ***, 24 unordered pieces.
ACCESSION AL390375
VERSION
KEYWORDS HTG: HTGS_PHASE1.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Burton, J.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On Aug 28, 2000 this sequence version replaced gi:9801494.
-----
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
Project Information
Center project name: BA505119
-----
Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 112223 bases at least Q40
Consensus quality: 119000 bases at least Q30
Consensus quality: 123719 bases at least Q20
Insert size: 131288; sum-of-contigs
Insert size: 148251; 6.5% error; agarose-fp
Quality coverage: 2.34x in Q20 bases; sum-of-contigs Quality
coverage: 2.28x in Q20 bases; agarose-fp
-----
```

\* NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 2237: contig of 2237 bp in length
2238 2337: gap of 100 bp
2338 9726: contig of 7389 bp in length
9727 9830: gap of 104 bp
9831 25051: contig of 15221 bp in length
25052 25151: gap of 100 bp
25152 27892: contig of 2741 bp in length
27893 27992: gap of 100 bp
27993 34854: contig of 6862 bp in length
34855 34954: gap of 100 bp
34955 37912: contig of 2958 bp in length
37913 38013: gap of 100 bp
38013 41365: contig of 3353 bp in length
41366 41465: gap of 100 bp
41466 45854: contig of 4389 bp in length
45855 45954: gap of 100 bp
45955 49072: contig of 3118 bp in length
49073 49172: gap of 100 bp
49173 52275: contig of 3103 bp in length
52276 52375: gap of 100 bp
52376 54684: contig of 2309 bp in length
54685 54784: gap of 100 bp
54785 59627: contig of 4843 bp in length
59628 59727: gap of 100 bp
59728 73263: contig of 13536 bp in length
73264 73363: gap of 100 bp
73364 76654: contig of 3291 bp in length
76655 76754: gap of 100 bp
76755 83253: contig of 6499 bp in length
83254 83353: gap of 100 bp
83354 86267: contig of 2914 bp in length
86268 86367: gap of 100 bp
86368 90797: contig of 4430 bp in length
90798 90897: gap of 100 bp
90898 96810: contig of 5913 bp in length
96811 96910: gap of 100 bp
96911 103462: contig of 6352 bp in length
103463 103562: gap of 100 bp
103563 111941: contig of 8379 bp in length
111942 112041: gap of 100 bp
112042 115654: contig of 3613 bp in length
115655 115754: gap of 100 bp
115755 121966: contig of 6212 bp in length
121967 122066: gap of 100 bp
122067 126044: contig of 3978 bp in length
126045 126144: gap of 100 bp
126145 133588: contig of 7444 bp in length.
```

## FEATURES

```
source
1. 133588
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-505I19"
/clone_1lb="RPC1-11.2"
1. 2237
/note="assembly_fragment:00181"
fragment_chain:1
2338..9726
/note="assembly_fragment:00638"
fragment_chain:1
9827..25051
/note="assembly_fragment:00200"
fragment_chain:2
25152..27892
/note="assembly_fragment:00042"
```

## misc\_feature

```
2338..9726
/note="assembly_fragment:00638"
fragment_chain:1
```

## misc\_feature

```
9827..25051
/note="assembly_fragment:00200"
fragment_chain:2
```

## misc\_feature

```
25152..27892
/note="assembly_fragment:00042"
```

```
misc_feature      fragment_chain:2"
27993. .34854
/note="assembly_fragment:00788
fragment_chain:3"
34955. .37912
/note="assembly_fragment:00826
fragment_chain:3"
38013. .41365
/note="assembly_fragment:01164
fragment_chain:4"
41466. .45854
/note="assembly_fragment:00675
fragment_chain:4"
45955. .49072
/note="assembly_fragment:00079"
49173. .52275
/note="assembly_fragment:00331"
52376. .54684
/note="assembly_fragment:00431"
54785. .59627
/note="assembly_fragment:00445"
59728. .73263
/note="assembly_fragment:00471"
73364. .76654
/note="assembly_fragment:00548"
76755. .83253
/note="assembly_fragment:00655"
83354. .86267
/note="assembly_fragment:00731"
86368. .90797
/note="assembly_fragment:00752"
90898. .96810
/note="assembly_fragment:00800"
96911. .103462
/note="assembly_fragment:00920"
103563. .111941
/note="assembly_fragment:00965"
112042. .115654
/note="assembly_fragment:01083"
115755. .121966
/note="assembly_fragment:01092"
122067. .126044
/note="assembly_fragment:01132"
126145. .133588
/note="assembly_fragment:01178"
2324 others

BASE COUNT 43601 a 22133 c 22728 g 42802 t
ORIGIN

Query Match      2.0%  Score 22;  DB 81;  Length 133588;
Best Local Similarity 100.0%;  Pred. No. 1.1;
Matches 22;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  635  TTGTTTCTTTATCCCTTGAT 656
      |||||
Db  8174  TTGTTTCTTTATCCCTTGAT 8153

RESULT 9
AL359812      AL359812 153203 bp  DNA  HTG  23-JAN-2001
LOCUS      Homo sapiens chromosome X clone RP11-488I3, *** SEQUENCING IN
DEFINITION  PROGRESS ***, 14 unordered pieces.
ACCESSION  AL359812
VERSION    AL359812.3  GI:10178576
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 153203)
AUTHORS    Pavitt,R.
TITLE      Direct Submission
```

## JOURNAL

## COMMENT

Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone  
requests: clonerequest@sanger.ac.uk  
On Sep 15, 2000 this sequence version replaced gi:9798060.  
----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: BA488I3  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: plasmid; L08752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 146625 bases at least Q40  
Consensus quality: 149438 bases at least Q30  
Consensus quality: 150830 bases at least Q20  
Insert size: 151903; sum-of-contigs  
Insert size: 160846; 5.7% error; agarose-fp  
Quality coverage: 4.24x in Q20 bases; sum-of-contigs Quality  
coverage: 4.04x in Q20 bases; agarose-fp  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 14 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 15556: contig of 15556 bp in length  
\* 15557 15656: gap of 100 bp  
\* 15657 34501: contig of 18845 bp in length  
\* 34502 34601: gap of 100 bp  
\* 34602 37279: contig of 2678 bp in length  
\* 37280 37379: gap of 100 bp  
\* 37380 49571: contig of 12192 bp in length  
\* 49572 49671: gap of 100 bp  
\* 49672 67178: contig of 17507 bp in length  
\* 67179 67278: gap of 100 bp  
\* 67279 74960: contig of 7682 bp in length  
\* 74961 75060: gap of 100 bp  
\* 75061 77751: contig of 2691 bp in length  
\* 77752 77851: gap of 100 bp  
\* 77852 81169: contig of 3318 bp in length  
\* 81170 81269: gap of 100 bp  
\* 81270 84632: contig of 3363 bp in length  
\* 84633 84732: gap of 100 bp  
\* 84733 102431: contig of 17699 bp in length  
\* 102432 102531: gap of 100 bp  
\* 102532 110974: contig of 8443 bp in length  
\* 110975 111074: gap of 100 bp  
\* 111075 116646: contig of 5572 bp in length  
\* 116647 116746: gap of 100 bp  
\* 116747 128736: contig of 11990 bp in length  
\* 128737 128836: gap of 100 bp  
\* 128837 153203: contig of 24367 bp in length.

## FEATURES

## source

1. .153203  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/clone\_lib="RPCI-11.2"  
1. .15556  
/note="assembly\_fragment:00087  
fragment\_chain:1"  
15657. .34501  
/note="assembly\_fragment:00981  
fragment\_chain:1"  
34602. .37279  
/note="assembly\_fragment:00571  
fragment\_chain:1"

## misc\_feature

15657. .34501

## misc\_feature

34602. .37279

## misc\_feature

15657. .34501

misc\_feature fragment\_chain:1"  
 37380..49571  
 /note="assembly\_fragment:00473  
 fragment\_chain:2"  
 49672..67178  
 /note="assembly\_fragment:00886  
 fragment\_chain:2"  
 67279..74960  
 /note="assembly\_fragment:00902  
 fragment\_chain:3"  
 75061..77751  
 /note="assembly\_fragment:01009  
 fragment\_chain:3"  
 77852..81169  
 /note="assembly\_fragment:00151"  
 81270..84632  
 /note="assembly\_fragment:00370"  
 84733..102431  
 /note="assembly\_fragment:00664"  
 102532..110974  
 /note="assembly\_fragment:00708"  
 111075..116646  
 /note="assembly\_fragment:00856"  
 116747..128736  
 /note="assembly\_fragment:01063"  
 128837..153203  
 /note="assembly\_fragment:00281  
 clone\_end:SP6  
 vector\_side:right"  
 49180 a 25365 c 25558 g 51792 t 1308 others

BASE COUNT 49180 a 25365 c 25558 g 51792 t 1308 others  
 ORIGIN

Query Match 2.0%; Score 22; DB 80; Length 153203;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 635 TTGGTTCTTTATCCCTTGTAT 656  
 |||||  
 Db 39947 TTGGTTCTTTATCCCTTGTAT 39968

RESULT 10  
 AL391873/c  
 LOCUS AL391873 161317 bp DNA HTG 07-FEB-2001  
 DEFINITION Homo sapiens chromosome 13 clone RP11-399N11, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, in unordered pieces.  
 ACCESSION AL391873  
 VERSION AL391873.13 GI:12718113  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 161317)  
 Almeida,J.  
 Direct Submission  
 TITLE Submitted (06-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 JOURNAL requests: clonerequest@sanger.ac.uk  
 COMMENT On Feb 8, 2001 this sequence version replaced gi:11876081.  
 ----- Genome Center  
 Center: Sanger Centre  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk  
 ----- Project Information  
 Center project name: BA399N11  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; L08752; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Consensus quality: 161302 bases at least Q40

Consensus quality: 161317 bases at least Q30  
 Consensus quality: 161317 bases at least Q20  
 Insert size: 161317; sum-of-contigs  
 Insert size: 159554; 2.2% error; agarose-fp  
 Quality coverage: 10.66x in Q20 bases; sum-of-contigs Quality  
 coverage: 11.10x in Q20 bases; agarose-fp  
 -----  
 \* NOTE: This is a 'working draft' sequence.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

#### FEATURES

source Location/Qualifiers

1..161317  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="13"  
 /clone="RP11-399N11"  
 /clone\_lib="RPC1-11.2"  
 1..161317  
 /note="assembly\_fragment:02944  
 clone\_end:SP6  
 vector\_side:left  
 clone\_end:T7  
 vector\_side:right"  
 misc\_feature 52121 a 28193 c 27654 g 53349 t

BASE COUNT  
 ORIGIN

Query Match 2.0%; Score 22; DB 81; Length 161317;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 635 TTGGTTCTTTATCCCTTGTAT 656  
 |||||  
 Db 42237 TTGGTTCTTTATCCCTTGTAT 42216

RESULT 11  
 AC068783/c

LOCUS AC068783 40020 bp DNA INV 12-MAY-2000  
 DEFINITION Leishmania major chromosome 35 clone L108 strain Friedlin, complete  
 sequence.  
 ACCESSION AC068783  
 VERSION AC068783.2 GI:7770358  
 KEYWORDS HTG.  
 SOURCE Leishmania major.  
 ORGANISM Leishmania major  
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
 Leishmania.  
 1 (bases 1 to 40020)  
 Myler,P.J., Sisk,E., Cawthra,J., Handley,F., Vogt,C., Robertson,L.,  
 McDonagh,P., Ivens,A., Nguyen,D., Munden,H. and Stuart,K.  
 Direct Submission  
 TITLE Submitted (09-MAY-2000) Seattle Biomedical Research Institution, 4  
 Nickerson Street, Seattle, WA 98109-1651, USA  
 JOURNAL 2 (bases 1 to 40020)  
 Myler,P.J., Sisk,E., Cawthra,J., Handley,F., Vogt,C., Robertson,L.,  
 McDonagh,P., Ivens,A., Nguyen,D., Munden,H. and Stuart,K.  
 Direct Submission  
 TITLE Submitted (12-MAY-2000) Seattle Biomedical Research Institution, 4  
 Nickerson Street, Seattle, WA 98109-1651, USA  
 JOURNAL On May 12, 2000 this sequence version replaced gi:7740033.  
 COMMENT NOTE: This sequence is still preliminary and may contain some  
 errors, such as substitutions and frameshifts. Please regard all  
 CDS sequences with some degree of caution. A finished version  
 should be available within a few weeks.

#### FEATURES

source Location/Qualifiers  
 1..40020  
 /organism="Leishmania major"  
 /strain="Friedlin"  
 /db\_xref="taxon:5664"  
 /chromosome="35"  
 /clone="L108"

```
gene      complement(432. .1580)
          /gene="L108.1"
          /note="predicted using Glimmer, Testcode and CodonUsage"
CDS       complement(432. .1580)
          /gene="L108.1"
          /note="Blastp similarity at N-terminus to SGT and
          hypothetical proteins from several organisms, also similar
          to LmS7il from Leishmania"
          /codon_start=1
          /product="L108.2"
          /protein_id="AAF69725.1"
          /db_xref="GI:7770359"
          /translation="MVNAFAEELKARGNEAFAAKNFEEAIVLYDKALEVDSNFIYNN
          NRAAAYHELNKAKIEDANKSISIENNAHARHLAGALWAOMKYRAAKNEFEVAATM
          DPKSSTIKDLEILNPMASTSAVANRGVPHPYEYARAAANAANAALVCGGEFV
          SVETGIGLVDVAVVLAALQIVASMLAESMASTLWAYVLLITMGQALVMVRNLL
          QFMKDILNSHFCSLFLTCFFAQLTRVRPIMFVVFIAAGVVDLVHKKRQQAAL
          EKVETFIAPYMOQASDAKAIIFITATVEALLLFTYMGFTGGFTFLVYIYQAYRYN
          HDGVYKLAFLKGLRLNITRMPTKSFMPQAVDRYAQKFFDLLDKVASQPI"
          complement(2750. .5752)
          /gene="L108.2"
          /note="predicted using Glimmer, Testcode and CodonUsage"
CDS       complement(2750. .5752)
          /gene="L108.2"
          /note="Blastp similarity to several proteins including
          plectin and myosin"
          /codon_start=1
          /product="L108.2"
          /protein_id="AAF69726.1"
          /db_xref="GI:7770360"
          /translation="MSAAMSRNYPASFAAKDDDKHYKMKCKIAQLTKVIVTLNRCED
          NEQRTWLHOTHMQRQAQCOEYSHLEKIQOOIRGLQADKINIAASLERSOKELOQA
          AQEEFTRLAALSEQSEQSRVDSSTILLEKQRMASLAEATKRVREGKDELANLYR
          EYNDKYVMLAEQDSLEALEKLRWEGSKVDALQRELAAGSAAQASQSKRLIQ
          EALQCATLANRDVSAASDEAHTQIQTLOGRTEELAGERRORAAQSDLEKTEL
          EARLSIDSKRNALQCEAEQAQSLAAEKALTARQAAYVALENEKASLMTSTALHQ
          ELGECQAALTATQADAAQRLDQSSSTHAILSDVRSQLEDEAKQRLATLHQEA
          AEARITSOAKHEELKARASLESNSSEQORLMEAHKKESDALKRTHEAELRLQTE
          MORLTERATNLSLORTAKGAEDAIAALTAQCEHEDALLRSRAEESAERQOLE
          QOLEQOREVEARKGCAEQAAQAALALERAQAEHAACAAEAHAHAKALREAGEKSA
          EQUESVRNNAALVAAQKQSAAMERAKVAAQAASISICRNAEAQAALRRDYDERL
          ASVYRAAATEGLHTLIQSELAQRLQALSETAALQTSNSDMLEQQAETALLKAAME
          KAQTEAAVLDTRSDAQDRDEHVALEVARQWAAEQRYEAAVAAAGKEQAKME
          ARDERIEAAKHLEQTAAALKAARIKDMQELEOLNMAHQLRVRELTATHDHAEKGL
          QSTIGTMDIKLELISMSAKEKLDRLLAQRLQLEEVATAKENGAAVVAEONRA
          ADLEQLRVEHORELKALGDTHKEIVDFVRQQAEEERRAHQGLAEELSTVEELQYKTH
          YRESRPEDVELINKLQDNKEKELALSAYEDMRKLYKLELINREENYKVFGRHPNVA
          DYSAVAQAAGKRPSSLPRIGATROHSTR"
          complement(7162. .8211)
          /gene="L108.3"
          /note="predicted using Glimmer, Testcode and CodonUsage"
CDS       complement(7162. .8211)
          /gene="L108.3"
          /codon_start=1
          /product="L108.3"
          /protein_id="AAF69727.1"
          /db_xref="GI:7770361"
          /translation="MLKWLURGSSSEAEKPAEKKPVSRMVVDAEKPLETHESAASRR
          REAYEKKLAESAKEGPPHPSKMITYDGFIRKAOQGVLLGNNNETQEGITLINARAQN
          AMISTKTLVNPOMSHVEWNLMNGFSDIIVAASWNTLNRYQLMYRVSSSTGAMLVTOF
          MAQKOGMSGOTVFAMTQYVWRFGGCTQVOYVKDOSLGLSHVORLIRGVHVTNLTV
          PTHNSYLSHAISLTTAKRDAQFMAEMTPSKGTWKLAAAFDMSMNDAAIELEYMEG
          REGKSNALNVCGRKSFVGGQAQLTSSLGFSMAKNVNDLDPFGGEVPGANQLRLTFNCQI
          DHAGALKOGLVFETA"
          complement(9855. .10607)
          /gene="26SPAS"
          /note="L108.4"
          /complement(9855. .10607)
          /gene="26SPAS"
          /note="component of 26S proteasome; Blastp similarity to
          alpha or iota subunit of 26S proteasome from several
          organisms"
          /codon_start=1
```

```
gene      complement(14214. .14984)
          /gene="L108.5"
          /note="predicted using Glimmer, Testcode and CodonUsage"
CDS       complement(14214. .14984)
          /gene="L108.5"
          /codon_start=1
          /product="L108.5"
          /protein_id="AAF69729.1"
          /db_xref="GI:7770363"
          /translation="MYASLTPTPPHLCPESSFFSLLSQVLSVASVAPGCBEFEFLS
          HPILPHLLLYFFQFOALATVICHCLWETLFLFAVFLRHLILPIRLPLSPVHFVGFV
          LILGFASCTYTHMGCVTARAARDVMANSKSKLSREESVQAAAEKPPSEAEVDYV
          AQFVAGYTDYDQKGSVTVYEDYRGFETSTTNHVSKNKADKCLKWSDQALEDLNLNN
          DVFNPQPFSPWPKEVITNHLSPSPFADRNVTQVEMTRK"
          complement(15912. .17309)
          /gene="L108.6"
          /note="predicted using Glimmer, Testcode and CodonUsage"
CDS       complement(15912. .17309)
          /gene="L108.6"
          /codon_start=1
          /product="L108.6"
          /protein_id="AAF69730.1"
          /db_xref="GI:7770364"
          /translation="MHRSSRVWGCSCGAPRTAARAFSGSVREDITEAERRHLNQELP
          PRELRELTDTATAGEDEDKIKVPGDSFGIMHNVFVKPQROPLSGRIGELMPL
          GTHVULCEAPDETVMRMSRIGVPVDVVSALRDSAGDNLATELLQRLDLIEVGFSGY
          GIVULCSTAPETTCILVSYNLPTASTQADEVLDAIHELTSAAELPLDTPDRDLVDKF
          TSHWTTEDVPCADMLRAYNITVQNILLPLFGDYGQGVHLLPKPKETNIGVAAA
          CMMDLRTGHNRFREHVIADSVSEHVQVMEHLHYNQGVHLLPQVWFRPEYVEEII
          RFKESLOPASTFEMRYAFLEAPQALPHRYNIVEMEKLKIAHQYKHYEDFTAPG
          KWLTSNDAQLOTVAAGGGSNVPVGGAMHNHANDPASIRTAMETRAGPLRRTLQASMQA
          HGRVFSRFYRNHH"
          complement(18724. .19731)
          /gene="L108.7"
          /complement(18724. .19731)
          /gene="L108.7"
          /codon_start=1
          /product="L108.7"
          /protein_id="AAF69731.1"
          /db_xref="GI:7770365"
          /translation="MFFWVVALLVIAIVMTVTPAGKLAQKQYAHGVVFMVKAEDL
          PKAYSGTNYIRMNTVGTGGKRVIRVIRHGSVYNSLNFNSCNLRPLVRAAKAAR
          EFTDEFTNPFGSCIIDSPSSKSKKREAOQLANFMRTAKTKISFDPTTSVVVSSNLRRA
          METALIGVSPRLTVTOERIVYMSSLSQEGSONIDAQSLSTEPGKIACKLTIIDPLEL
          AVFDPYLNAGNRVAGVDVYQRMDEFIMHLFGGSGEANLYPAAGGSNAALKEVIVVGH
          SCYFRNFRFLPHSTHISKKRKMQNCVAVAFELTRAETSGEVAVEESTLKVLYKGF
          A"
          complement(22260. .23417)
          /gene="L108.8"
          /note="predicted using Glimmer, Testcode and CodonUsage"
CDS       complement(22260. .23417)
          /gene="L108.8"
          /codon_start=1
          /product="L108.8"
          /protein_id="AAF69732.1"
          /db_xref="GI:7770366"
          /translation="MTAPGIWACNSCSVEQMABEATYCMCKFARPMDRRGVPPQIFSQY
          KIHFNGLIIPRTIVHPSHSVSWERHAGATCCTSFDPVAVSVILVRYGYSKRCRIC
          LEQHLNIPCVPIAWLLDSLOSROIHPSLYLRLLPLVANPTVGTGDLPHOHPFYOI
          NKFEYSITPSFASKSKASKSAATGDGANROVEYSGEIEAALPFFDIEPRCTAMSV
          FDAVACTGTGKTEADADDITNIEARKTPGIELLAQAQASNKVDRALFSGMNVLLIP
          SLQQTAVVMVTIQRCGKKAERKRESLEATLRSGVTIYIYSHEDKDKDEVMIOAAHLST
          SLPGQLQVSNWLEDCLILGELLPLRGMVTPPTAKLTITLNKKHTKAKSCAV"
          complement(24639. .29171)
          /gene="AMPD2"
```

```

/note="L108.9; predicted using Glimmer, Testcode and
CodonUsage"
complement(24639. .29171)
/genre="AMPD2"
/note="Involved in purine metabolism: Blastp similarity to
AMP deaminases from several organisms, including
Leishmania (L302.11) with which it is not identical"
/codon_start=1
/product="AMP deaminase 2"
/protein_id="AAF69733.1"
/db_xref="GI:7770367"
/translation="MFLASCSFORLPLPHLFTFTFTSLVTMPLSISMMDSKSN
VGDSTAMPPELYHAGLLOPASPHNSASTAFSEKRGOLHSCVPHEDLTREPE
MFQVYDGDQVEMRRVHDIETALRVSLYRPLETRVGGRRANPNYSAMPGRIT
IQVKGVTQVSDHDSLEFLPIWQIATDVOKRLTVGNAGCVNACHRLGLIMQERS
RFFFLNAGMSTRELHKGAVRSADLTAVAGLQWAEKADPAHQGLDLACDSMAALG
AVHLRNGSNTRELHKGAVRSADLTAVAGLQWAEKADPAHQGLDLACDSMAALG
AELRESFTEQLCEKLVLRVVSRAERSLTPQAAEYSPVLYQSSLSYLAELMO
RLLEGPHRVQYILSTCTFSPFVSVSSCTTLOQDLNIFLALFKATLAPEDPSNAG
VANLLGQVGGLOMLHAQDQGRDFEDMAPPPQVQKAGQSLGYMYIYANLAVLNS

Query Match      1.9%; Score 21; DB 4; Length 40020;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 TCATAGCTGTGGTCTGTGT 760
      |||||
Db 6058 TCATAGCTGTGGTCTGTGT 6038

RESULT 12
AL157780/c
LOCUS AL157780 156623 bp DNA HTG 23-JAN-2001
DEFINITION Homo sapiens chromosome 9 clone RP11-98E22 map q33.1-33.3, ***
SEQUENCING IN PROGRESS ***, 16 unordered pieces.
ACCESSION AL157780
VERSION AL157780.4 GI:9212428
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 156623)
Plumb,B.
Direct Submission
Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 15, 2000 this sequence version replaced gi:8051973.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA98E22
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 14635 bases at least Q40
Consensus quality: 14908 bases at least Q30
Consensus quality: 15208 bases at least Q20
Insert size: 155123; sum-of-contigs
Insert size: 171845; 3.3% error; agarose-fp
Quality coverage: 3.42x in Q20 bases; sum-of-contigs Quality
coverage: 3.09x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```

```

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 5730: contig of 5730 bp in length
5731 5830: gap of 100 bp
5831 7902: contig of 2072 bp in length
7903 8002: gap of 100 bp
8003 10546: contig of 2544 bp in length
10547 10646: gap of 100 bp
10647 35183: contig of 24537 bp in length
35184 35283: gap of 100 bp
35284 55290: contig of 20007 bp in length
55291 55390: gap of 100 bp
55391 62727: contig of 7337 bp in length
62728 62827: gap of 100 bp
62828 65872: contig of 3045 bp in length
65873 65972: gap of 100 bp
65973 81607: contig of 15635 bp in length
81608 81707: gap of 100 bp
81708 96593: contig of 14886 bp in length
96594 96693: gap of 100 bp
96694 103342: contig of 6649 bp in length
103343 103442: gap of 100 bp
103443 116947: contig of 13505 bp in length
116948 117047: gap of 100 bp
117048 119556: contig of 2509 bp in length
119557 119656: gap of 100 bp
119657 127113: contig of 7457 bp in length
127114 127213: gap of 100 bp
127214 146933: contig of 19720 bp in length
146934 147033: gap of 100 bp
147034 151142: contig of 4109 bp in length
151143 151242: gap of 100 bp
151243 156623: contig of 5381 bp in length.
Location/Qualifiers
1..156623
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="q33.1-33.3"
/clone="RP11-98E22"
/clone_lib="RPCI-11.1"
1..5730
/note="assembly_fragment:00931
fragment_chain:1
clone_end:SP6
vector_side:left"
5831..7902
/note="assembly_fragment:01181
fragment_chain:1"
8003..10546
/note="assembly_fragment:01024
fragment_chain:2"
10647..35183
/note="assembly_fragment:00119
fragment_chain:2"
35284..55290
/note="assembly_fragment:00552
fragment_chain:2"
55391..62727
/note="assembly_fragment:00333
fragment_chain:3"
62828..65872
/note="assembly_fragment:00497
fragment_chain:3"
65973..81607
/note="assembly_fragment:01267
fragment_chain:4"
81708..96593
/note="assembly_fragment:00449
fragment_chain:4"
96694..103342
/note="assembly_fragment:00001"

```

```

misc_feature      103443..116947
/note="assembly_fragment:00089"
misc_feature      117048..119556
/note="assembly_fragment:00149"
misc_feature      119657..127113
/note="assembly_fragment:00713"
misc_feature      127214..146933
/note="assembly_fragment:01087"
misc_feature      147034..151142
/note="assembly_fragment:01118"
misc_feature      151243..156623
/note="assembly_fragment:00696
clone_end:T7
vector_side:right"
BASE COUNT      48164 a 28785 c 28083 g 50089 t 1502 others
ORIGIN.

```

```

Query Match      1.98; Score 21; DB 79; Length 156623;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 533 AAAAATACACACCCAGGCA 553
|||||
Db 17748 AAAAATACACACCCAGGCA 17728

```

```

RESULT 13
AP002432/c
LOCUS      AP002432      139515 bp      DNA      HPG      06-JUN-2000
DEFINITION Homo sapiens chromosome 11 clone RP11-113J20 map 11q14, WORKING
DRAFT SEQUENCE, 14 unordered pieces.
ACCESSION      AP002432
VERSION        AP002432.1 GI-8307736
KEYWORDS       HG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 139515)
AUTHORS        Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE          Homo sapiens 139,515 genomic DNA of 11q14
JOURNAL        Published Only in DataBase (2000) In press
REFERENCE      2 (bases 1 to 139515)
AUTHORS        Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE          Direct Submission
JOURNAL        Submitted (02-JUN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
COMMENT        ----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP11-113J20
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator EF-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 128683 bases at least Q40
Consensus quality: 134465 bases at least Q30
Consensus quality: 136763 bases at least Q20
Insert size: 138215; sum-of-contigs
Quality coverage: 5.3lx in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of

```

```

14 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 27253 contig of 27253 bp in length
27354 44336 contig of 16983 bp in length
44437 61944 contig of 17508 bp in length
62045 77522 contig of 15478 bp in length
77623 87441 contig of 9819 bp in length
87542 96973 contig of 9432 bp in length
97074 104498 contig of 7425 bp in length
104599 111686 contig of 7088 bp in length
111787 120891 contig of 9105 bp in length
120992 126947 contig of 5956 bp in length
127048 132775 contig of 5728 bp in length
132876 137159 contig of 4284 bp in length
137260 138741 contig of 1482 bp in length
138842 139515 contig of 674 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 27253: contig of 27253 bp in length
27254 27353: gap of 100 bp
27354 44336: contig of 16983 bp in length
44337 44437: gap of 100 bp
44437 61944: contig of 17508 bp in length
61945 62044: gap of 100 bp
62045 77522: contig of 15478 bp in length
77523 77622: gap of 100 bp
77623 87441: contig of 9819 bp in length
87442 87541: gap of 100 bp
87542 96973: contig of 9432 bp in length
96974 97073: gap of 100 bp
97074 104498: contig of 7425 bp in length
104499 104598: gap of 100 bp
104599 111686: contig of 7088 bp in length
111687 111786: gap of 100 bp
111787 120891: contig of 9105 bp in length
120892 120991: gap of 100 bp
120992 126947: contig of 5956 bp in length
126948 127047: gap of 100 bp
127048 132775: contig of 5728 bp in length
132776 132875: gap of 100 bp
132876 137159: contig of 4284 bp in length
137160 137259: gap of 100 bp
137260 138741: contig of 1482 bp in length
138742 138841: gap of 100 bp
138842 139515: contig of 674 bp in length.
FEATURES
Location/Qualifiers
1. 139515
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q14"
/clone="RP11-113J20"
1. 27253
/note="assembly_fragment"
27354..44336
/note="assembly_fragment"
44437..61944
/note="assembly_fragment"
62045..77522
/note="assembly_fragment"
77623..87441
/note="assembly_fragment"
87542..96973
misc_feature
1. 27253
misc_feature
27354..44336
misc_feature
44437..61944
misc_feature
62045..77522
misc_feature
77623..87441
misc_feature
87542..96973

```

```
misc_feature /note="assembly_fragment"
97074..104498
/note="assembly_fragment clone_end:SP6 vector_side:right"
104599..111686
/note="assembly_fragment"
111787..120891
/note="assembly_fragment"
120992..126947
/note="assembly_fragment"
127048..132775
/note="assembly_fragment"
132876..137159
/note="assembly_fragment"
137260..138741
/note="assembly_fragment"
138842..139515
/note="assembly_fragment clone_end:T7 vector_side:left"
41925 a 26963 c 27931 g 41396 t 1300 others

BASE COUNT 41925 a 26963 c 27931 g 41396 t 1300 others
ORIGIN

Query Match 1.8%; Score 20; DB 83; Length 139515;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 TCACAGAGTCCTGGCTTC 891
|||||
Db 8983 TCACAGAGTCCTGGCTTC 8964

RESULT 14
AP002796
LOCUS AP002796 158155 bp DNA HTG 03-NOV-2000
DEFINITION Homo sapiens chromosome 11 clone RP11-742N3 map 11g, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
ACCESSION AP002796
VERSION AP002796.2 GI:11094149
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-742N3.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 158155)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 158,155 genomic DNA of 11g
Published Only in DataBase (2000) In press
2 (bases 1 to 158155)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (11-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel.81-45-503-9111, Fax.81-45-503-9170)
On Nov 3, 2000 this sequence version replaced gi:9188582.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@psc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP11-742N3
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 155377 bases at least Q40
Consensus quality: 156614 bases at least Q30
Consensus quality: 157219 bases at least Q20
Insert size: 157455; sum-of-contigs
```

```
Quality coverage: 10.45x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
8 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 38031 contig of 38031 bp in length
38132 contig of 26303 bp in length
64535 contig of 25159 bp in length
89794 contig of 27617 bp in length
117511 contig of 17509 bp in length
135120 contig of 9634 bp in length
144854 contig of 7757 bp in length
152711 contig of 5445 bp in length.
NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 38031: contig of 38031 bp in length
38032 38131: gap of 100 bp
38132 64434: contig of 26303 bp in length
64435 64534: gap of 100 bp
64535 89693: contig of 25159 bp in length
89694 89793: gap of 100 bp
89794 117410: contig of 27617 bp in length
117411 117510: gap of 100 bp
117511 135019: contig of 17509 bp in length
135020 135119: gap of 100 bp
135120 144753: contig of 9634 bp in length
144754 144853: gap of 100 bp
144854 152610: contig of 7757 bp in length
152611 152710: gap of 100 bp
152711 158155: contig of 5445 bp in length.
Location/Qualifiers
1..158155
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11g"
/clone="RP11-742N3"
1..38031
/note="assembly_fragment"
38132..64434
/note="assembly_fragment"
64535..89693
/note="assembly_fragment"
89794..117410
/note="assembly_fragment"
117511..135019
/note="assembly_fragment"
135120..144753
/note="assembly_fragment"
144854..152610
/note="assembly_fragment clone_end:SP6 vector_side:right"
152711..158155
/note="assembly_fragment clone_end:T7 vector_side:left"
BASE COUNT 46482 a 30851 c 31484 g 48638 t 700 others
ORIGIN

Query Match 1.8%; Score 20; DB 83; Length 158155;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 TCACAGAGTCCTGGCTTC 891
|||||
```

Db 18247 TCACAGAAGTCCTGGCTTTC 18266

## RESULT 15

AC025876/c

## LOCUS

AC025876 159804 bp DNA HTG 26-MAY-2000  
Homo sapiens clone RP11-537L13, WORKING DRAFT SEQUENCE, 14  
unordered pieces.

## ACCESSION

AC025876

## VERSION

AC025876.4 GI:8077111

## KEYWORDS

HTG; HTGS\_PHASE1; HTGS\_DRAFT.

## SOURCE

Human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 159804)

## AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

## TITLE

Unpublished

## REFERENCE

2 (bases 1 to 159804)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavskiy, L., Boukhaltier, B., Brown, A., Burkett, G.,  
Campopiano, A., Castile, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Glinde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., Lakocek, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marguis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Willson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

## DIRECT SUBMISSION

Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 25, 2000 this sequence version replaced gi:7923927.

## COMMENT

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

## ----- Project Information

Center project name: L8150  
Center clone name: 537\_L13  
----- Summary StatisticsSequencing vector: MJ3; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152340 bases at least Q40  
Consensus quality: 156099 bases at least Q30  
Consensus quality: 157400 bases at least Q20Insert size: 157000; agarose-fp  
Insert size: 158504; sum-of-contigs  
Quality coverage: 4.6 in Q20 bases; agarose-fp  
Quality coverage: 4.5 in Q20 bases; sum-of-contigs----- NOTE: This is a 'working draft' sequence. It currently  
consists of 14 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.\* 1 475: contig of 475 bp in length  
\* 476 575: gap of 100 bp  
\* 576 2356: contig of 1781 bp in length  
\* 2357 2456: gap of 100 bp  
\* 2457 5028: contig of 2572 bp in length  
\* 5029 5128: gap of 100 bp  
\* 5129 7548: contig of 2420 bp in length  
\* 7549 7648: gap of 100 bp  
\* 7649 11846: contig of 4198 bp in length  
\* 11847 18125: contig of 6179 bp in length  
\* 18126 18225: gap of 100 bp  
\* 18226 26080: contig of 7855 bp in length  
\* 26081 26180: gap of 100 bp  
\* 26181 34006: contig of 7826 bp in length  
\* 34007 34106: gap of 100 bp  
\* 34107 47967: contig of 13861 bp in length  
\* 47968 48067: gap of 100 bp  
\* 48068 65206: contig of 17139 bp in length  
\* 65207 65306: gap of 100 bp  
\* 65307 85368: contig of 20262 bp in length  
\* 85369 85668: gap of 100 bp  
\* 85669 107683: contig of 22015 bp in length  
\* 107684 107783: gap of 100 bp  
\* 107784 131012: contig of 23229 bp in length  
\* 131013 131112: gap of 100 bp  
\* 131113 159804: contig of 28692 bp in length.

## FEATURES

## source

1. 159804  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-537L13"  
/clone\_lib="RPC1-11 Human Male BAC"

## misc\_feature

1. 475  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right

## misc\_feature

576..2356  
/note="assembly\_fragment"  
2457..5028  
/note="assembly\_fragment"

## misc\_feature

5129..7548  
/note="assembly\_fragment"  
clone\_end:Y7  
vector\_side:right

## misc\_feature

7649..11846  
/note="assembly\_fragment"  
11947..18125  
/note="assembly\_fragment"

## misc\_feature

18226..26080  
/note="assembly\_fragment"  
26181..34006  
/note="assembly\_fragment"

## misc\_feature

34107..47967  
/note="assembly\_fragment"  
48068..65206  
/note="assembly\_fragment"

## misc\_feature

65307..85368  
/note="assembly\_fragment"  
85669..107683  
/note="assembly\_fragment"

## misc\_feature

107784..131012  
/note="assembly\_fragment"  
131113..159804  
/note="assembly\_fragment"BASE COUNT 48298 a 31570 c 31477 g 47156 t 1303 others  
ORIGIN

Query Match

1.8%; Score 20; DB 70; Length 159804;

```

Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 TCACAGAACTCTCGGTTTC 891
|||||
Db 24260 TCACAGAACTCTCGGTTTC 24241

RESULT 16
AL354868/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-339A7 on chromosome 6, complete
sequence.
ACCESSION AL354868
VERSION AL354868.10 GI:10862757
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164018)
Direct Submission
Submitted (16-OCT-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Oct 17, 2000 this sequence version replaced gi:10715936.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/Celegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-339A7 is from the library RPCI-11.2 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-339A7 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-339A7 is at 164018 in this
sequence. The true right end of clone RP3-486D24 is at 100 in this
sequence.

FEATURES
source
1..164018
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone_lib="RP11-339A7"
/clone_lib="RPCI-11.2"
9..125
repeat_region
/note="L1M4 repeat: matches 3373..3490 of consensus"
126..483
repeat_region
/note="MER7A repeat: matches 1..346 of consensus"
484..608
repeat_region
/note="L1M4 repeat: matches 3490..3611 of consensus"
619..1243
repeat_region
/note="L1P repeat: matches 8..639 of consensus"

```

```

1244..1906
/note="L1PA4 repeat: matches 5483..6146 of consensus"
1918..2683
/note="L1M4 repeat: matches 3608..4379 of consensus"
2684..2982
/note="AluSx repeat: matches 6..304 of consensus"
2983..3084
/note="L1M4 repeat: matches 4379..4459 of consensus"
3075..3374
/note="AluY repeat: matches 1..301 of consensus"
3552..3791
/note="HERVL40 repeat: matches 5278..5522 of consensus"
3856..4260
/note="L1R40a repeat: matches 79..518 of consensus"
4680..4721
/note="21 copies 2 mer ac 100% conserved"
4924..5351
/note="L1PA5 repeat: matches 5708..6143 of consensus"
5482..5646
/note="AluSg/x repeat: matches 130..294 of consensus"
9041..9325
/note="AluSq repeat: matches 1..305 of consensus"
9668..10320
/note="L1MC1 repeat: matches 5661..6319 of consensus"
10306..11117
/note="L1ME repeat: matches 5184..5390 of consensus"
11118..11395
/note="AluSx repeat: matches 1..292 of consensus"
11396..11933
/note="L1ME repeat: matches 5390..5947 of consensus"
12298..12601
/note="MT1B repeat: matches 163..466 of consensus"
12602..12710
/note="L1M1 repeat: matches 6189..6304 of consensus"
12711..12798
/note="MT1B repeat: matches 76..163 of consensus"
12799..13172
/note="MER47A repeat: matches 1..366 of consensus"
13173..13249
/note="MT1B repeat: matches 1..77 of consensus"
13966..14057
/note="HAL1 repeat: matches 688..775 of consensus"
14515..14826
/note="match: GSS: Em:B15656"
15424..15652
/note="MIR repeat: matches 27..261 of consensus"
15641..15749
/note="MER5B repeat: matches 12..104 of consensus"
19358..19573
/note="MER58A repeat: matches 2..224 of consensus"
complement(20165..20463)
/note="match: GSS: Em:B33496"
20493..20895
/note="match: GSS: Em:AQ340676"
20579..20897
/note="match: GSS: Em:AQ066536"
20587..20925
/note="match: GSS: Em:AQ067144"
20824..21138
/note="MER58B repeat: matches 1..341 of consensus"
21248..21556
/note="AluSx repeat: matches 7..312 of consensus"
21726..21892
/note="MIR repeat: matches 6..161 of consensus"
22216..22527
/note="AluSx repeat: matches 1..312 of consensus"
23412..23880
/note="L1P1 repeat: matches 5674..6153 of consensus"
23881..24186
/note="AluSx repeat: matches 1..299 of consensus"
24187..24204
/note="L1P11 repeat: matches 5658..5674 of consensus"
24345..24493

```

repeat\_region /note="L2 repeat: matches 2605. .2750 of consensus"  
24648. .24920  
repeat\_region /note="L1MC5 repeat: matches 7609. .7907 of consensus"  
25046. .25234  
repeat\_region /note="L2 repeat: matches 2546. .2748 of consensus"  
25301. .25355  
/note="MER4-internal repeat: matches 1112. .1162 of  
consensus"  
repeat\_region 25660. .25705  
misc\_feature /note="L2 repeat: matches 2661. .2706 of consensus"  
25724. .26160  
repeat\_region /note="match: GSS: Em:AQ617525"  
26864. .27231  
misc\_feature /note="MT1A2 repeat: matches 6. .370 of consensus"  
complement(26912. .27451)  
repeat\_region /note="match: GSS: Em:AQ442658"  
27338. .27926  
repeat\_region /note="L2 repeat: matches 2148. .2750 of consensus"  
28118. .28559  
repeat\_region /note="L2 repeat: matches 2273. .2703 of consensus"  
28833. .28874  
repeat\_region /note="MT1A2 repeat: matches 33. .73 of consensus"  
32887. .33067  
repeat\_region /note="MER57B repeat: matches 1. .403 of consensus"  
33513. .33558  
repeat\_region /note="L2 repeat: matches 2374. .2419 of consensus"  
34684. .35033  
repeat\_region /note="AluX repeat: matches 127. .309 of consensus"  
35034. .35077  
repeat\_region /note="22 copies 2 mer gt 95% conserved"  
35079. .35206  
repeat\_region /note="AluX repeat: matches 1. .129 of consensus"  
35240. .35373  
repeat\_region /note="AluJb repeat: matches 5. .134 of consensus"  
36077. .36399  
repeat\_region /note="AluJo repeat: matches 3. .302 of consensus"  
36728. .36793  
repeat\_region /note="33 copies 2 mer ta 69% conserved"  
36935. .36992  
repeat\_region /note="29 copies 2 mer at 70% conserved"  
37000. .37227  
repeat\_region /note="MT1D repeat: matches 2. .248 of consensus"  
37228. .37613  
repeat\_region /note="THE1C repeat: matches 1. .371 of consensus"  
37614. .37689  
repeat\_region /note="MT1D repeat: matches 248. .319 of consensus"  
37771. .37955  
repeat\_region /note="MT1D repeat: matches 315. .504 of consensus"  
38215. .39023  
repeat\_region /note="pTR5 repeat: matches 762. .1519 of consensus"  
39385. .39800  
misc\_feature /note="Cpg Island"  
/evidence=not\_experimental  
repeat\_region 39024. .39661  
repeat\_region /note="LTR12 repeat: matches 1. .671 of consensus"  
41270. .41550  
repeat\_region /note="AluSg repeat: matches 24. .302 of consensus"  
41569. .41767  
repeat\_region /note="L1MC5 repeat: matches 7715. .7929 of consensus"  
42607. .42660  
misc\_feature /note="27 copies 2 mer ac 88% conserved"  
complement(43439. .43946)  
repeat\_region /note="match: GSS: Em:AQ764042"  
43454. .43519

Query Match 1.8%; Score 20; DB 89; Length 164018;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 293 CCACTGGTGGTGGTTC 312  
|||||  
Db 75240 CCACTGGTGGTGGTTC 75221

RESULT 17

AC024192

LOCUS

DEFINITION

AC024192

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC024192 164907 bp DNA HTG 07-JUL-2000  
Homo sapiens chromosome 4 clone RP11-476H13, WORKING DRAFT  
SEQUENCE, 19 unordered pieces.  
AC024192  
HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 164907)  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 164907)  
Waterston,R.H.  
Direct Submission  
Submitted (25-FEB-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On May 25, 2000 this sequence version replaced gi:7523979.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

----- Project Information -----

Center project name: H.NH0476H13

----- Summary Statistics -----

Sequencing vector: M13; 100%

Sequencing method: plasmid; 0%

Chemistry: Dye-primer ET; 79% of reads

Chemistry: Dye-terminator Big Dye; 21% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 156472 bases at least Q40

Consensus quality: 158945 bases at least Q30

Consensus quality: 160470 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 163107; sum-of-contigs

Quality coverage: 5.04 in Q20 bases; agarose-fp

Quality coverage: 5.29 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently

\* consists of 19 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1264: contig of 1264 bp in length

\* 1265 1364: gap of unknown length

\* 1365 3092: contig of 1728 bp in length

\* 3093 3193: gap of unknown length

\* 3193 4314: contig of 1122 bp in length

\* 4315 4415: gap of unknown length

\* 4415 5875: contig of 1361 bp in length

\* 5876 9008: gap of unknown length

\* 9009 9108: contig of 3133 bp in length

\* 9109 12786: contig of 3678 bp in length

\* 12787 12886: gap of unknown length

\* 12887 16247: contig of 3361 bp in length

\* 16248 16347: gap of unknown length

\* 16348 20262: contig of 3915 bp in length

\* 20263 20363: gap of unknown length

\* 20363 24659: contig of 4297 bp in length

\* 24660 24759: gap of unknown length

\* 24760 29152: contig of 4393 bp in length

\* 29153 29252: gap of unknown length

```

* 29253 33764: contig of 4512 bp in length
* 33765 33864: gap of unknown length
* 33865 40006: contig of 6142 bp in length
* 40007 40106: gap of unknown length
* 40107 46699: contig of 6593 bp in length
* 46700 46799: gap of unknown length
* 46800 54943: contig of 8144 bp in length
* 54944 55043: gap of unknown length
* 55044 64712: contig of 9669 bp in length
* 64713 64812: gap of unknown length
* 64813 82487: contig of 17675 bp in length
* 82488 82587: gap of unknown length
* 82588 103480: contig of 20893 bp in length
* 103481 103580: gap of unknown length
* 103581 133552: contig of 29972 bp in length
* 133553 133652: gap of unknown length
* 133653 164907: contig of 31255 bp in length.
FEATURES
    source
        Location/Qualifiers
            1..164907
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="4"
            /clone="RP11-476H13"
BASE COUNT 50190 a 30972 c 30730 g 51211 t 1804 others
ORIGIN
Query Match 1.8%; Score 20; DB 69; Length 164907;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 657 GTTCATGATATTTGTACA 676
|||||
Db 6158 GTTCATGATATTTGTACA 6157

RESULT 18
AP000793 AP000793 169152 bp DNA PRI 23-JAN-2001
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-812L3.
ACCESSION AP000793
VERSION AP000793.4 GI:12381913
KEYWORDS
SOURCE
    ORGANISM
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    REFERENCE
        1 (bases 1 to 169152)
        Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
        Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
        Homo sapiens 169,152 genomic DNA of 11q
        Published Only in Database (1999) in press
    REFERENCE
        2 (bases 1 to 169152)
        Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
        Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
        Direct Submission
        Submitted (29-NOV-1999) Masahira Hattori, The Institute of Physical
        and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
        1-7-22 Sushiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
        (E-mail:hattori@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
        Tel:81-45-503-9111, Fax:81-45-503-9170)
    COMMENT
        On Jan 22, 2001 this sequence version replaced gi:9927298.
FEATURES
    source
        Location/Qualifiers
            1..169152
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11q"
            /clone="RP11-812L3"
BASE COUNT 49748 a 31219 c 32888 g 55297 t
ORIGIN

```

```

Query Match 1.8%; Score 20; DB 90; Length 169152;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 872 TCACAGAAGTCCTGGCTTTC 891
|||||
Db 17969 TCACAGAAGTCCTGGCTTTC 17988
|||||
RESULT 19
AC010903 AC010903 179401 bp DNA HTG 14-DEC-2000
LOCUS Homo sapiens chromosome 2 clone RP11-550H9, WORKING DRAFT SEQUENCE,
DEFINITION 3 unordered pieces.
ACCESSION AC010903
VERSION AC010903.7 GI:11465184
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
    ORGANISM
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    REFERENCE
        1 (bases 1 to 179401)
        Waterston,R.H.
        The sequence of Homo sapiens clone
        Unpublished
    REFERENCE
        2 (bases 1 to 179401)
        Waterston,R.H.
        Direct Submission
        Submitted (25-SEP-1999) Genome Sequencing Center, Washington
        University School of Medicine, 4444 Forest Park Parkway, St. Louis,
        MO 63108, USA
    COMMENT
        On Nov 30, 2000 this sequence version replaced gi:11323453.
        ----- Genome Center -----
        Center: Washington University Genome Sequencing Center
        Center code: WUGSC
        Web site:http://genome.wustl.edu/gsc/index.shtml
        ----- Project Information -----
        Center project name: H_NH0550H09
        ----- Summary Statistics -----
        Sequencing vector: M13; 49%
        Chemistry: Dye-primer ET; 37% of reads
        Chemistry: Dye-terminator Big Dye; 63% of reads
        Assembly program: Phrap; version 0.990319
        Consensus quality: 176429 bases at least Q40
        Consensus quality: 177554 bases at least Q30
        Consensus quality: 178057 bases at least Q20
        Insert size: 174000; agarose-fp
        Insert size: 179201; sum-of-contigs
        Quality coverage: 6.28 in Q20 bases; agarose-fp
        Quality coverage: 6.19 in Q20 bases; sum-of-contigs
        -----
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 3 contigs. The true order of the pieces
        * is not known and their order in this sequence record is
        * arbitrary. Gaps between the contigs are represented as
        * runs of N, but the exact sizes of the gaps are unknown.
        * This record will be updated with the finished sequence
        * as soon as it is available and the accession number will
        * be preserved.
        * 1 34792: contig of 34792 bp in length
        * 34793 34892: gap of unknown length
        * 34893 75248: contig of 40356 bp in length
        * 75249 75348: gap of unknown length
        * 75349 179401: contig of 104053 bp in length.
FEATURES
    source
        Location/Qualifiers
            1..179401
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="2"
            /clone="RP11-550H9"
            misc_feature 1..34792

```

```
/note="assembly_name:Contig11"
34893. .75248
/note="assembly_name:Contig12"
75349. .179401
/note="assembly_name:Contig13
clone_end:SP6
vector_side:right"
BASE COUNT 58216 a 32501 c 32026 g 56457 t 201 others
ORIGIN

Query Match 1.8%; Score 20; DB 62; Length 179401;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 AGTCTCTAAAGGCATCTAT 348
|||||
DB 136764 AGTCTCTAAAGGCATCTAT 136783

RESULT 20
AC079533 201978 bp DNA HTG 02-SEP-2000
LOCUS Mus musculus clone RP23-348M1, WORKING DRAFT SEQUENCE, 7 ordered
DEFINITION pieces.
AC079533
AC079533 1 GI:9964898
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 201978)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Mouse
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 201978)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1870793
Center clone name: RPCI-23_348M1
-----
Summary Statistics
Consensus quality: 197680 bases at least Q40
Consensus quality: 200626 bases at least Q30
Consensus quality: 201212 bases at least Q20
Estimated insert size: 205000; agarose-fp estimation
Estimated insert size: 201728; sum-of-contigs estimation
Quality coverage: 10.36 in Q20 bases; agarose-fp estimation
Quality coverage: 10.53 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 14235: contig of 14235 bp in length
* 14236 14335: gap of unknown length
* 14336 40527: contig of 26192 bp in length
* 40528 40627: gap of unknown length
* 40628 49494: contig of 8867 bp in length
* 49495 49594: gap of unknown length
```

```
* 49595 89951: contig of 40357 bp in length
* 89952 90051: gap of unknown length
* 90052 152065: contig of 62014 bp in length
* 152066 152165: gap of unknown length
* 152166 193308: contig of 41143 bp in length
* 193309 193409: gap of unknown length
* 193409 201978: contig of 8570 bp in length.
FEATURES
Location/Qualifiers
1..201978
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RPCI mouse BAC library 23"
BASE COUNT 55239 a 43169 c 44386 g 58583 t 601 others
ORIGIN

Query Match 1.8%; Score 20; DB 76; Length 201978;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 879 AGTCCTGGCTTCTCTCACT 898
|||||
DB 109347 AGTCCTGGCTTCTCTCACT 109366

RESULT 21
G39250 494 bp DNA STS 01-FEB-2001
LOCUS Z20386 zebrafish AB Danio rerio STS genomic, sequence tagged site.
DEFINITION G39250
ACCESSION G39250
VERSION G39250.1 GI:3358459
KEYWORDS STS.
SOURCE zebrafish.
ORGANISM Danio rerio
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Rasbora; Danio.
1 (bases 1 to 494)
Shimoda, N., Knapik, E.W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S.,
Jackson, D., de Sauvage, F., Jacob, H. and Fishman, M.C.
Zebrafish genetic map with 2000 microsatellite markers
Genomics 58 (3), 219-232 (1999)
99303552
COMMENT
Contact: Mark C. Fishman
Cardiovascular Research Center
Massachusetts General Hospital
Mail code 1494100A, 149 13th Street, Charlestown, MA 02129, USA
Fax: 6177265806
Email: fishman@mgh.cvr.harvard.edu
http://zebrafish.mgh.harvard.edu
Primer A: AATCCTCATCTCTCTGCATG
Primer B: TGCAGACTCAGAAAAGCATC
STS size: 255
PCR Profile:
Presoak: 94 degrees C for 5.0 minutes
Denaturation: 94 degrees C for 1.0 minute
Annealing: 58 degrees C for 1.0 minute
Polymerization: 72 degrees C for 1.5 minute
PCR Cycles: 27
Thermal Cycler: MJ Research PTC-100
Protocol:
Template: 10 ng
each 375 nM
dNTPs: each 200 uM
Tag Polymerase: 0.034 units/ul
Total Vol: 10 ul
Buffer:
MgCL2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
```

PH: 8.3

Primers are available from Research Genetics Inc.  
(http://www.resgen.com phone: 800-533-4363).

FEATURES  
source  
1. .494  
/organism="Danio rerio"  
/strain="AB"  
/db\_xref="taxon:7955"  
/clone\_lib="zebrafish AB"  
/sex="F"  
/dev\_stage="Adult"  
/lab\_host="DH5alphaF'IQ"  
/note="Vector: ml3mp19 with added BstXI site; V-type:  
Phage: Genomic DNA from a single adult zebrafish of AB  
strain was digested with AluI, Cae8I, HaeIII, NlaVI, or  
RsaI. Fragments in the range of 250-500 bp were gel  
purified and a BstXI linker was added. The fragments were  
cloned into a modified M13mp19 vector and transformed  
into E. Coli DH5alpha. Microsatellite sequences were  
screened with labeled d(CA)15 and d(GT)15 oligonucleotide  
probes."  
STS  
primer\_bind 99. .353  
primer\_bind 99. .118  
BASE COUNT 129 a 108 c 117 g 139 t 1 others  
ORIGIN  
Query Match 1.7%; Score 19; DB 54; Length 494;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 ACTTTAACTGCGGATGCT 374  
|||||  
Db 321 ACTTTAACTGCGGATGCT 339  
|||||

RESULT 22  
LOCUS CNS06EZW 840 bp DNA STS 10-JAN-2001  
DEFINITION T3 end of clone AR0AA026A11 of library AR0AA from strain CBS 732 of  
Zygosaccharomyces rouxii, sequence tagged site.  
ACCESSION AL395778  
VERSION AL395778.1 GI:12147596  
KEYWORDS STS.  
SOURCE Zygosaccharomyces rouxii.  
ORGANISM Zygosaccharomyces rouxii  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.  
REFERENCE 1 (bases 1 to 840)  
AUTHORS de Montigny, J., Straub, M., Potier, S., Tekai, F., Dujon, B.,  
Wincker, P., Artiguenave, F. and Souciet, J.  
TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 8.  
JOURNAL Zygosaccharomyces rouxii(1)  
PUBMED FEBS Lett. 487 (1), 52-55 (2000)  
REFERENCE 2 (bases 1 to 840)  
AUTHORS Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,  
Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,  
de Montigny, J., Dujon, B., Durand, P., Lepingle, A., Liorente, B.,  
Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S.,  
Saurin, W., Tekai, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,  
Wincker, P. and Weissenbach, J.  
TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of  
yeast species for molecular evolution studies(1)  
JOURNAL FEBS Lett. 487 (1), 3-12 (2000).  
PUBMED 11152876  
REFERENCE 3 (bases 1 to 840)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :

seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
This STS is part of a random genomic sequencing program of thirteen  
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,  
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,  
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of  
the other extremity of this insert.

FEATURES  
Location/Qualifiers  
source  
1. .840  
/organism="Zygosaccharomyces rouxii"  
/strain="CBS 732"  
/db\_xref="taxon:4956"  
/clone\_lib="AR0AA026A11"  
/clone\_lib="AR0AA"  
/note="end : T3"  
STS  
1. .840  
BASE COUNT 240 a 168 c 174 g 256 t 2 others  
ORIGIN

Query Match 1.7%; Score 19; DB 53; Length 840;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 AGTCATATTCATATTAC 53  
|||||  
Db 778 AGTCATATTCATATTAC 796  
|||||

RESULT 23  
LOCUS AF072760/c 1989 bp mRNA ROD 23-JUL-1998  
DEFINITION Mus musculus fatty acid transport protein 5 mRNA, complete cds.  
ACCESSION AF072760  
VERSION AF072760.1 GI:3335570  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 1989)  
AUTHORS Hirsch, D., Stahl, A. and Lodish, H.F.  
TITLE A family of fatty acid transporters conserved from mycobacterium to  
man  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (15), 8625-8629 (1998)  
MEDLINE 98337965  
REFERENCE 2 (bases 1 to 1989)  
AUTHORS Stahl, A., Hirsch, D. and Lodish, H.F.  
TITLE Direct Submission  
JOURNAL Submitted (21-JUN-1998) Whitehead Institute, 9 Cambridge Center,  
Cambridge, MA 02142, USA

FEATURES  
Location/Qualifiers  
source  
1. .1989  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
CDS  
1. .1989  
/note="FATP5"  
/codon\_start=1  
/product="fatty acid transport protein 5"  
/protein\_id="AAC40189.1"  
/db\_xref="GI:3335571"  
/translation="MALALRWFLGDPFCVLVLLGLLGRPWISSWMPHWLSLVCAALT  
LFLPLQPPGLRWLHKDAFTFMFLFYGLKFRRLNKKHPETFDALERQALAWPOR  
VALVCTGSEGSITNSQDARSQAAVWLKAKLDKDAVQTRDAATLVLPSTISAL  
SVFLGLAKLQCPVAINPHSRGMPPLHSRVSASGLIVDPDQENLEEVLPKLLAEN  
IHCYFLGHSPTGVEALGASDAAPSDVPASLRATIKWSPAIFFTFSGTGLPKP  
AIIASHRVQVSNVLFCGCRADDVVDVLPVHTIGLVGLGCLQVGATCVLAPKF  
SASRFWAECRQHGVTYILYGEILRYLCNVYPEQEDKIHVRLAMGTGLRANWKNFQ  
QRGPRIRIWEFYGTSGTGNVGLMNVVGHGAVGRTSCILRMLTFPELVQFDIETAEPLR

DKOGFCIPVEGRKPGLLTKVRKNQPELGYRGSQAESNRKLVANVRVGDLYFNTGDV  
 LTLQDEFFYQDRLGDTFRWKGENVSTGEVECVLSLDLEEVNVTGVPVPGCEGKV  
 GMAVKLAPKGTGDKGLQYHRSWLPAYATPHFIRIODSLSEINTYKLVKSRVREG  
 FDVGLIADPLYILDNAQTSRLMPDQVYQAVCEGTWNL"

BASE COUNT 425 a 525 c 547 g 492 t  
 ORIGIN

Query Match 1.7%; Score 19; DB 94; Length 1989;

Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 876 AGAAGTCTGCTTCCTG 894

Db 1400 AGAAGTCTGCTTCCTG 1382

RESULT 24

AB038484 2072 bp DNA PLN 15-NOV-2000  
 LOCUS  
 DEFINITION Caulerpa okamurae chloroplast rbcL gene for  
 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit,  
 partial cds.

ACCESSION AB038484

VERSION AB038484.1 GI:11182373

KEYWORDS ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit.  
 Caulerpa okamurae chloroplast DNA.

SOURCE Chloroplast Caulerpa okamurae  
 ORGANISM Eukaryota; Viridiplantae; Chlorophyta; Ulvophyceae; Caulerpaceae;  
 Caulerpaceae; Caulerpa.

REFERENCE 1 (sites)

AUTHORS Hanyuda,T., Arai,S. and Ueda,K.

TITLE Variability in the rbcL Introns of Caulerpaean Algae (Chloroph  
 yta, Ulvophyceae)

JOURNAL J. Plant Res. (2000) In press

REFERENCE 2 (bases 1 to 2072)

AUTHORS Hanyuda,T. and Ueda,K.

TITLE Direct Submission

JOURNAL Submitted (17-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takeaki  
 Hanyuda, Kanazawa University, Graduate School of Natural Science  
 and Technology, Kakuma, Kanazawa, Ishikawa 920-1192, Japan  
 (E-mail:hanyu@kenroku.kanazawa-u.ac.jp, Tel:+81-76-264-5706,  
 Fax:+81-76-264-5976)

COMMENT Sequence updated (11-Mar-2000).

FEATURES Location/Qualifiers

source

1..2072

/organism="Caulerpa okamurae"

/organelle="plastid:chloroplast"

/db\_xref="taxon:118247"

join(1..220,960..2072)

/gene="rbcL"

join(<1..220,960..2072)

/gene="rbcL"

/codon\_start=2

/product="ribulose-1,5-bisphosphate carboxylase/oxygenase

large subunit"

/protein\_id="BAB17910.1"

/db\_xref="GI:11182374"

/translation="GAGFAGVKDYRLTYTPKVKLETDILAAFRMTPOGVPPEEA

GAVAAESSTGTTVTWDTGLTSLDKYKRCYDIEAVTGEENQYIAYVAPLDLFEEG

SVNPLFTSVGNVFGFKALRALLEDLRIPPAYAKTFQGPFGHIGVERDKLNFYGRSL

LGCTIRPKGLSKATNGRAVYECILRGDLGDTKDDENVNSQPFMRWRDRFLFVAEAIYK

AOATEGEIKGHYNATAGTCDDEMIKRAQCAKDFGVPITIMHDYLTGGFTANTSLAFGR

DHGLLLIHRAMHADVIRDKIHGMHFRVLAKALRLSGDHLHSGTVYVKGLEGEVFL

GFVLMRDIFVEKDRSRGVYFTQDASLPGVIVPASGGIHWHPALVEIFGDDCLQ

FGGGTLGHFWGNAPGAANRIACEIQARNEGRSLAAGNQIIRDAARNSP"

221..959

intron

BASE COUNT 571 a 329 c 471 g 701 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 19; DB 12; Length 2072;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 167 TTCTGGTGGTGATCACCTT 185  
 Db 1673 TTCTGGTGGTGATCACCTT 1691

RESULT 25

AB013985/c 2248 bp DNA PLN 06-APR-1999  
 LOCUS  
 DEFINITION Antirrhinum majus transposon Tam3 pseudogene for transposase (in  
 S-5 copy).

ACCESSION AB013985

VERSION AB013985.1 GI:3219242

KEYWORDS transposase.

SOURCE Antirrhinum majus DNA.

ORGANISM Antirrhinum majus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Asteridae; euasterids I; Lamiales; Scrophulariaceae;  
 Antirrhinum.

REFERENCE 1 (sites)

AUTHORS Kishima,Y., Yamashita,S., Martin,C. and Mikami,T.

TITLE Structural conservation of the transposon Tam3 family in

Antirrhinum majus and estimation of the number of copies able to

transpose

JOURNAL Plant Mol. Biol. 39 (2), 299-308 (1999)

MEDLINE 99178781

AUTHORS Kishima,Y.

REFERENCE 2 (bases 1 to 2248)

TITLE Direct Submission

JOURNAL Submitted (14-MAY-1998) to the DDBJ/EMBL/GenBank databases. Yuji  
 Kishima, Faculty of Agriculture, Hokkaido University, Lab. Genetic  
 Engineering, Kita 9, Nishi 9, Kita-ku, Sapporo, Hokkaido 060-8589,  
 Japan (E-mail:geiabeabs.agr.hokudai.ac.jp, Tel:011-706-2484,  
 Fax:011-716-0879)

FEATURES Location/Qualifiers

source

1..2248

/organism="Antirrhinum majus"

/transposon="Tam3"

/db\_xref="taxon:4151"

/note="S-5 copy"

1..>2248

/codon\_start=1

/pseudo

/product="transposase"

BASE COUNT 756 a 379 c 448 g 665 t

ORIGIN

Query Match

Best Local Similarity 1.7%; Score 19; DB 12; Length 2248;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 530 ACCAAATACACACCCA 548

Db 77 ACCAAATACACACCCA 59

RESULT 26

MMAJ3959/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

MMAJ3959 2310 bp mRNA ROD 22-JUL-1998

Mus musculus mRNA for very-long-chain acyl-CoA synthetase related

protein (VLACSR).

Accession

AJ223959

very-long-chain acyl-CoA synthetase related protein; vlacsr gene.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2310)

Berger,J.

TITLE Direct Submission  
JOURNAL Submitted (29-JAN-1998) Berger J., Neuropathology and  
Neurochemistry, Institute of Neurology, University of Vienna,  
Schwarzschanenstr. 17, A-1090 Vienna, AUSTRIA

REMARK Revised by author 20-APR-98

REFERENCE 2 (bases 1 to 2310)  
AUTHORS Berger J., Truppe C., Neumann H. and Forss-Petter, S.  
TITLE A novel relative of the very-long-chain acyl-CoA synthetase and  
fatty acid transporter protein genes with a distinct expression  
pattern

JOURNAL Biochem. Biophys. Res. Commun. 247 (2), 255-260 (1998)  
MEDLINE 98308102

FEATURES Location/Qualifiers  
1..2310  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/sex="Male"  
/dev\_stage="11 weeks"  
/tissue\_type="whole liver"  
89..2158  
/gene="vlacsr"  
89..2158  
/gene="vlacsr"  
/note="strong expression in liver, low expression in lung,  
brain, testis, spleen and skeletal muscle, no expression  
in kidney and heart"  
/codon\_start=1  
/product="very-long-chain acyl-CoA synthetase related  
protein"  
/protein\_id="CAAL1688.1"  
/db\_xref="GI:3341462"  
/translation="MGWKLTLLLLLLVGLGPPWPAAMALRLWFLGPTCLVL  
LGLALGRPWISSWPHWLSVGAALTLFLPLPPGLRWLHIDVAFTKMLFYGLK  
FRRLNKHPTTFVDALEROALAWDRVALVCTSEGSSITNSOLDARSQAAWLKA  
KLKDAVIONTDAAILVPSKTSIALSVFLGLAKLGPVANINPHSRGMPLLHSVRS  
SGASVLIIDPDQLENEVLKLAENHCFYLGHSPTPGVEALGASLDAAPSDVP  
ASLRATKGPSPAPIFFTSGTGLPKPAILSHERVIOVNSLFCGRADDDVYDLP  
LYHTIGLVGLQGVATCVLAPKFSARFACRQHGVTIVLYGVGLRLYCNVP  
EQPEDKHTVRLAMGNGLRANVKNFQRFGRIRIWEFGYSGEYVGLMNYVGHGAV  
GRTECILRLMPLFELQVDFIETAPLRQKQFCIPVEPKGKGLLLTKVRKNQPFILGYR  
GSQESLFLFENVYGVPGCEGKGVMAAVKAPGFTFGCKLQHVRSWLPATAT  
ESVLSSDLFENVYGVPGCEGKGVMAAVKAPGFTFGCKLQHVRSWLPATAT  
PHEFTQLSEITNTYIKLVKSLRVGDFVGIADPLYLIDNKAOTFRSLMFDVQVAV  
CEGWTKL"  
polya\_signal 2264..2269  
polya\_site 2280  
BASE COUNT 535 a 594 c 620 g 561 t  
ORIGIN

Query Match 1.7%; Score 19; DB 94; Length 2310;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 876 AGAAGTCCTGCTTCTCCTG 894  
|||||  
Db 1569 AGAAGTCCTGCTTCTCCTG 1551

RESULT 27  
AK026973  
LOCUS Homo sapiens CDNA: FLJ23320 fis, clone HEP12381.  
DEFINITION AK026973  
ACCESSION AK026973  
VERSION AK026973.1 GI:10439963  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens hepatoma cell\_line:HepG2 cDNA to mRNA, clone\_lib:HEP  
clone:HEP12381.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)  
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,  
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,  
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished (2000)  
REFERENCE 2 (bases 1 to 2337)  
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,  
Shibahara,T., Tanaka,T. and Nakamura,Y.  
TITLE Direct Submission  
JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio  
Sugano, Institute of Medical Science, University of Tokyo,  
Laboratory of Genome Structure Analysis, Human Genome Center,  
Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
(E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
Fax:81-3-5449-5416)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing: Research Association for Biotechnology; cDNA library  
construction, 5'- & 3'-end one pass sequencing: Department of  
Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency).

FEATURES Location/Qualifiers  
1..2337  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="HepG2"  
/cell\_type="hepatoma"  
/clone="HEP12381"  
/clone\_lib="HEP"  
/note="cloning vector pME18SFL3"  
117..2195  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BABI5609.1"  
/db\_xref="GI:10439964"  
/translation="MKQFACLYLCSKVDYVRKILKPHSSILRTWLSKCQSPGPG  
NSNIFLQHRVENDQLYQYCSLLISIPKQLOMDPSSHQFGMDFGLGLDAD  
ETPLASTVLLMAVGIFGHWRTPLGYFVNRSAGYLQAQLRLTIGLSDITVLAV  
TSDATSVQMAKALGTHIDGDMKCTFOHPSSSQOIAVFFDSCHLLRLRNAFONF  
QSQIOLINGIAHWOHLVELVALEQELSNMERIPSTLANLKNHLKVNATOLFESVA  
SALGYLLSLDLPFQNCIGTIFHLRLINLFDIENSRCYCKGLKGPLLPETYSKINH  
VLEAKTIFVLTSDSNQIIGKQKGLGFLNLAESLAWLIQNYVFPKVPPEYLL  
TYFSDHDLFLKMLRQVLVTSSTPMAPQKAYNLETRYKDFODEVFLSKVFDI  
SIARRKDLALWTQRYGVSTVTFHEEGICQDWSHCSLEALLDLSHRNLICVA  
GYVANKLSALLTCEDCITALLYASDLKASKIGSLFVKKNGLHFPSESLCRVINICER  
VVRTHSRMAIFELVSKORELYLOQKILCELSGHIDLFDVNNKHLFDEGCVAINEFVKL  
LKDIICFLNIRAKNVAQNPLKHSERTDMKTLRKHWSSVDYKCSFANTSSKFRH  
LLSNDGTFPK"

BASE COUNT 765 a 399 c 440 g 733 t  
ORIGIN

Query Match 1.7%; Score 19; DB 89; Length 2337;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 943 AGAACTACTTCTTCTGAAGA 961  
|||||  
Db 13 AGAACTACTTCTTCTGAAGA 31

RESULT 28  
AF242189/c  
LOCUS AF242189 2360 bp mRNA ROD 02-SEP-2000  
DEFINITION Rattus norvegicus bile acid CoA ligase mRNA, complete cds.  
ACCESSION AF242189  
VERSION AF242189.1 GI:9963929  
KEYWORDS  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 2360)  
Falany,C.N., Xie,X., Wheeler,J., Wang,J. and Barnes,S.  
Molecular cloning and expression of rat liver bile acid CoA ligase  
Unpublished  
2 (bases 1 to 2360)  
Falany,C.N., Xie,X., Wheeler,J., Wang,J. and Barnes,S.  
Direct Submission  
Submitted (06-MAR-2000) Pharmacology & Toxicology, University of  
Alabama at Birmingham, 1670 University Blvd., Birmingham, AL 35294,  
USA

FEATURES	source
Location/Qualifiers	
1. .2360	/organism="Rattus norvegicus"
	/strain="Sprague-Dawley"
	/db_xref="taxon:10116"
52. .2124	
CDS	
	/codon_start=1
	/product="bile acid CoA ligase"
	/protein_id="AA09770.1"
	/db_xref="GI:9963930"
	/translation="MGVKKRLFTLLSLLLLLVLGQLMPAATAALRALRFLGDPTCFV LLGLAFLPRTSSIPHWLSAAALTLSLPPRPPELRWLHKDVAFAFKLLFTGL NLKRLEATRPPELFVDALBQQQAQRPDQVALVCTGSEGSITNRELNAKACQAAWLK AKLKREATIQEDKGATAILVQPSKATLSALSVFLGKLQCPVAINPHSRGMPLLHSVQ SSGVALVDPDQLEBVLKLLAENRCFLVHSITPTGVEALGADDAAPSDPV PAKLRAITKWSQAFIYTSQGTLPKPAILSHERIVSNVLSCPGRTADDAVYNV PLYHSMGLVLYGCLQGCATCLAPKFSASRYWAECROYSVTVILYGEVLYRLCNY PQQPCKKHTRFALGNLRADWENFQRPQPIQIWELYGSTEGNVNGLVYGHGCA VKGTSCFIRMLPTLELVQDITAEAPRQDQFCIPVETGKPEGLLTIKRKNQPFLLG RGSDETIRMPVRQGDLYNYNTKGVLDALDQEGFFYDRDLGDTFRKNGSLPVEA VEGVSILDLFLEVNYYGVTPGCEKVGMAAVKLAPGTFDQCKLYQHVRSMLWPYA TPHTRIQDLSLEITNTYKLVKSQALAREGFDVGVIADPLYILDNKAETFRSLMPDVYQA VCEGTWKL"

```

BASE COUNT      509 a      596 c      618 g      537 t
ORIGIN

Query Match      1.7%;      Score 19;      DB 94;      Length 2360;
Best Local Similarity 100.0%;      Pred. NO. 98;
Matches 19;      Conservative 0;      Mismatches 0;      Indels 0;      Gaps 0;

```

Qy	876	AGAACTCCTGGCTTTCCTG	894
D <sub>b</sub>	1535	AGAACTCCTGGCTTTCCTG	1517

RESULT	29
LOCUS	II2527
DEFINITION	Sequence 1 from patent US 5426041.
ACCESSION	II2527
VERSION	II2527.1 GI:909911
KEYWORDS	. Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 3293)
AUTHORS	Fabijanski,S.F. and Arinson,P.G.
TITLE	Binary cryptocytotoxic method of hybrid seed production
JOURNAL	Patent: US 5426041-A 1 20-JUN-1995;
FEATURES	Location/Qualifiers 1..3293 /organism='unknown'
source	
BASE COUNT	1032 a 701 c 579 g 968 t 13 others
ORIGIN	

```
Query Match      1.7%; Score 19; DB 10; Length 3293;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1097 CATGAAAGCTGAGTCTCC 1115  
|||||  
Db 2811 GATGAAAGCTGAGTCTCC 2829

RESULT 30  
HSBTDSS2 12990 bp DNA PRI 12-JUN-1998  
LOCUS Homo sapiens biotinidase (BTD) gene, exons 2, 3, and 4 and complete cds.  
DEFINITION HSBTDSS2 12990 bp DNA PRI 12-JUN-1998  
ACCESSION AF018631 GI:2674073  
VERSION AF018631.1  
KEYWORDS  
SEGMENT  
SOURCE  
ORGANISM Homo sapiens  
human.  
2 of 2  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 12990)  
AUTHORS Knight,H.C., Reynolds,T.R., Meyers,G.A., Pomponio,R.J., Buck,G.A.  
and Wolf,B.  
TITLE Structure of the human biotinidase gene  
JOURNAL Mamm. Genome 9 (4), 327-330 (1998)  
MEDLINE 98191738  
REFERENCE 2 (bases 1 to 12990)  
AUTHORS Knight,H.C. and Wolf,B.  
TITLE Direct Submission  
JOURNAL Submitted (12-AUG-1997) Human Genetics, Medical College of Virginia, 1101 East Marshall Street, Sanger Hall Room 11-002,  
Richmond, VA 23298, USA  
FEATURES  
source location/Qualifiers  
1..12990  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p25"  
order(AF018630.1:1..1000,1..12406)  
/gene="Btd"  
join(AF018630.1:566..644,3558..3822,10044..10193,  
10905..12406)  
/gene="Btd"  
/product="biotinidase"  
join(AF018630.1:601..644,3558..3822,10044..10193,  
10905..12077)  
/gene="Btd"  
/EC\_number="3.5.1.12"  
/function="hydrolyzes biocytin to the vitamin biotin and lysine"  
/note="biotin-amide amidohydrolase"  
/codon\_start=1  
/product="biotinidase"  
/protein\_id="AAC21679.1"  
/db\_xref="GI:2674075"  
/translation="MAHAHIQGRRARRKRVVCMISGASKLALFLCGCVVVALGHART  
KDEIVAFEDDIGHGFNFTSIYPFLDFMPSPQQVVRWNPCLPEHRFNDETVQLRSLC  
MALRGDMFTVANLGTEPKCHSDPRCPKDGRYOENTNVVFSNNGLVDRLVKKNLYFE  
AARDVPLKYDLTFDTPEAGRFGITCFDIILFDPAIRLVLDYKKVVVPTAMNQII  
PLLAIEIQKAFARVAGIIVLANVHHVPVLGMTCSGIHTPLESFVHDMDENPKSHLI I  
AQVAKPMVILGAENATGDDPSHKLILKSGDPYCEKDAQEHVCDEATKWVNAPPH  
THSEMMYNDFTLTPVMWGEGYLHVCSNGLCYLLFERPTLSKELYALGVFDGLHVVH  
GTVYIOCVALLRCGGIGFDTCGOETIATGIFEFHLWGNFSTSYIPFLFTSGMTILEY  
PDOLGWENDHYFLRKSRSLSSGLVTAAALYGLRYERD"  
<1..3557  
/gene="Btd"  
/number=1  
misc\_binding 3275..3284  
/gene="Btd"  
/bound\_moiety="C/EBP"  
misc\_binding 3520..3526  
/gene="Btd"  
/bound\_moiety="C/EBP"  
misc\_binding 3538..3546

/gene="Btd"  
 /bound\_moiety="Ap1"  
 3558..3822  
 /gene="Btd"  
 /number=2  
 3823..10043  
 /gene="Btd"  
 /number=2  
 10044..10193  
 /gene="Btd"  
 /number=3  
 10194..10904  
 /gene="Btd"  
 /number=3  
 10905..12406  
 /gene="Btd"  
 /number=4  
 12078..12406  
 /gene="Btd"  
 /note="3'UTR"  
 12374..12382  
 /gene="Btd"  
 BASE COUNT 3463 a 2950 c 2844 g 3733 t  
 ORIGIN

Query Match 1.7%; Score 19; DB 92; Length 12990;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 807 TGCTAATTGGGTAAATG 825  
 ||||||||||||||||  
 Db 1035 TGCTAATTGGGTAAATG 1053

RESULT 31  
 AC015050 28968 bp DNA HTG 16-NOV-1999  
 LOCUS Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered  
 DEFINITION pieces.  
 ACCESSION AC015050  
 VERSION AC015050.1 GI:6436285  
 KEYWORDS HTG: HTGS\_PHASE2.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 28968)  
 Adams, M. and Venter, J.C.  
 Direct Submission  
 TITLE Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,  
 JOURNAL Rockville, MD, USA  
 COMMENT This sequence was identified as CDM:10211102 by the submitter.  
 For further information on this sequence e-mail to fly@celera.com.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

FEATURES  
 source  
 1..28968  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 BASE COUNT 7995 a 6839 c 6833 g 7301 t  
 ORIGIN

Query Match 1.7%; Score 19; DB 64; Length 28968;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1034 AAAACATTTCTCGGAC 1052  
 ||||||||||||||||

Db 20124 AAAACATTTCTCGGACAC 20142

RESULT 32

AC020299 31857 bp DNA HTG 03-JAN-2000  
 LOCUS Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered  
 DEFINITION pieces.  
 ACCESSION AC020299  
 VERSION AC020299.1 GI:6664598  
 KEYWORDS HTG: HTGS\_PHASE2.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 31857)  
 Adams, M. and Venter, J.C.  
 Direct Submission  
 TITLE Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,  
 JOURNAL Rockville, MD, USA  
 COMMENT This sequence was identified as CDM:10212697 by the submitter.  
 For more information on this record e-mail to fly@celera.com.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

FEATURES  
 source  
 1..31857  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 BASE COUNT 8538 a 7036 c 7054 g 9229 t  
 ORIGIN

Query Match 1.7%; Score 19; DB 66; Length 31857;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 716 AAAGGCACAAAGGCATCCG 734  
 ||||||||||||||||  
 Db 26863 AAAGGCACAAAGGCATCCG 26881

RESULT 33

CEZK678 35126 bp DNA INV 27-OCT-2000  
 LOCUS Caenorhabditis elegans cosmid 2K678, complete sequence.  
 DEFINITION  
 ACCESSION Z79605  
 VERSION Z79605.1 GI:1515170  
 KEYWORDS HTG: Cyclin like; LIN-15A protein; Steroid hormone receptor.  
 SOURCE Caenorhabditis elegans.  
 ORGANISM Caenorhabditis elegans  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 35126)  
 none.  
 REFERENCE  
 AUTHORS Genome sequence of the nematode C. elegans: a platform for  
 TITLE investigating biology. The C. elegans Sequencing Consortium  
 JOURNAL Science 282 (5396), 2012-2018 (1998)  
 MEDLINE 99069613  
 REMARK The C. elegans Sequencing Consortium.  
 Erratum: [[published errata appear in Science 1999 Jan  
 1;283(5398):35 and 1999 Mar 26;283(5410):2103 and 1999 Sep  
 3;285(5433):1493]]  
 2 (bases 1 to 35126)  
 Kershaw, J.K.

REFERENCE  
 AUTHORS Direct Submission  
 TITLE Submitted (29-AUG-1996) Nematode Sequencing Project, Sanger Centre,  
 JOURNAL Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,  
 Washington University, St. Louis, MO 63110, USA. E-mail:  
 jes@sanger.ac.uk or rw@nematode.wustl.edu  
 COMMENT Coding sequences below are predicted from computer analysis, using

predictions from GeneFinder (P. Green, U. Washington), and other available information.  
 Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.  
 This sequence is the entire insert of clone ZK678. The true right end of clone ZK662 is at 2701 in this sequence. The start of this sequence (1..105) overlaps with the end of sequence Z79604. The end of this sequence (35024..35126) overlaps with the start of sequence AL034489.  
 For a graphical representation of this sequence and its analysis see: - <http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=ZK678>.

FEATURES	source	Location/Qualifiers	
		1..35126	
		/organism="Caenorhabditis elegans"	
		/db_xref="taxon:6239"	
		/chromosome="X"	
		/clone="ZK678"	
		join(484...4762,5230..5319,5371..5808,6314..6418,6466..6633,6692..7771)	
		/gene="lin-15"	
		/note="Identity to C.elegans LIN-15A protein (TR:G516019)"	
		CDNA EST Yk117a4.3 comes from this gene	
		CDNA EST Yk117a4.5 comes from this gene	
		CDNA EST Yk584d8.3 comes from this gene	
		CDNA EST Yk584d8.5 comes from this gene	
		/codon_start=1	
		/protein_id="CAB01904.1"	
		/db_xref="GI:3881845"	
		/db_xref="SPTREMBL:Q27365"	
		/translation="MLAPAPAKDVVSADKEEIIAKRFRPMKNVDMRWSSLANDRM AFNKKNALAMFKVSGAGTQDALQTCFQELVHFNPVIAAVVGVKREPNSVQAEEK KIPVKYITQTPTQSSVRLQOEKASATEQSAESASIMKHAFANTIPNSTQSVK DVTAAAGQKQSAEIEFISHPSEPSKPRATRECSQPSDYTYCTYLPCLCEKA LLMRESIANTDNEAVLMAAVMSGHFRMATAEKAIRHERLMCYDHVDVYEMWCDA FEAKTSEINEMPPDLMRGHDIYRALKRVGDLHGKVTNTPLYSEKNSIKSYRHH VPMVNGSLSKSPKPFSELVALLQSPVSTNLNHLNLSLDADKQBELIQLINGK DNFTSRRKIEDILDNKFAAAAKAYRDHSDAPSEYIPNQSEMQNTVYERRKKLH SEQDDAGSSISWNAKKTPTIDYVHLATRVLEHSIADEALLHKSVSYARNAFGE KFSPTPPAPLKFVVGNGKQYLFENGTEGPPKVVQGVNVLRTNLKDALYAPRAQ NPSTSTDSSEMSIEKQSFCAPOKEEERELVPTLQNKPTHVESSSPVEKKPPT KTNVEKPAVLGRMLTAFEGSMYSYRTKSVENKTDLLNQTSASPRMIKVVNRNPH LAKQVAAASEKPHIIPTHWEKKPEELMDPKPEPIF"	
		complement(join(11996..12256,12311..12526,12947..13135))	
		/gene="ZK678.6"	
		complement(join(11996..12256,12311..12526,12947..13135))	
		/codon_start=1	
		/protein_id="CAB01907.1"	
		/db_xref="GI:3881848"	
		/db_xref="SPTREMBL:O62517"	
		/translation="MSTSTVNLVNTLTSTTEILTYIOLTYGIPSPFFMMIFTLFLIA FGKVKSSPYTLVFDLSNVCVLTWAIETRMHPNVLVLLKAVEYTYIPGSLTWFK YPYWPFHMFWSALLTVHRLSSIFLPHRYESFNSRWFIILLIPIAICISHLPKYL WGTFLYEVIIIDVFIICFIHFLAKAAYNVAVFSVIYFLLNLSMGLITAFMVSKKFQ GNN"	
		14559..14642	
		/gene="ZK678.t1"	
		/note="CAG Leu L-trna"	
		predicted using tRNAscan-SE-1.11	
		preliminary prediction	
		similar to tRNA-Leu	
		14559..14642	
		/gene="ZK678.t1"	
		join(15119..15187,15733..16438,16512..16765)	
		/gene="ZK678.2"	
		join(15119..15187,15733..16438,16512..16765)	
		/gene="ZK678.2"	
		/note="predicted using GeneFinder	
		Similarity to steroid hormone receptors, contains	

similarity to Pfam domain: PF00104 (Ligand-binding domain of nuclear hormone receptor), Score=145.3, E-value=3.4e-40, N=1; PF00105 (Zinc finger, C4 type (two domains)), Score=32.8, E-value=4.9e-09, N=1"  
 /codon\_start=1  
 /protein\_id="CAB01903.1"  
 /db\_xref="GI:3881844"  
 /db\_xref="SPTREMBL:Q94411"  
 /translation="MTHFRKVENCIICDRETSLFNYGAISNACKLFFRRKAVSKTEI SPCSNRLCHTKSTWKSACKSCLLKCVKAGMTLNSTDONFMKTSRDYIDLQNL HLNLNTRMLASSHCVSQTDISIYIEIAQPSKISVPVPEALEMLNDWIFLOINTTI DFVKLPVAKDLNLSDRALFLKHQCFTLISLSYEAMSKKCYMSPDGSNLIPIQP SLKLENFANVKSQLVGRIQIRVTKKEYLLMVLFAFNPTDINTLSLYGKTLINTQOK EFSNALLYCFSNYQRTGSPSRGDLCLCSAVIKTRQDCSFFLLTNYPGPTFKLFNE IGQSFQNG"  
 complement(20315..20378)  
 /gene="ZK678.t5"  
 /note="TGG Pro P-trna"  
 predicted using tRNAscan-SE-1.11  
 preliminary prediction  
 similar to tRNA-Pro"  
 complement(20315..20378)  
 /gene="ZK678.t5"  
 20540..20610  
 /gene="ZK678.t2"  
 /note="TGG Pro P-trna"  
 predicted using tRNAscan-SE-1.11  
 preliminary prediction  
 similar to tRNA-Pro"  
 20540..20610  
 /gene="ZK678.t2"  
 complement(20825..20896)  
 /gene="ZK678.t4"  
 /note="TGG Pro P-trna"  
 predicted using tRNAscan-SE-1.11  
 preliminary prediction  
 similar to tRNA-Pro"  
 complement(join(20992..21064,21164..21274,21355..21500,21554..21715))  
 /gene="ZK678.3"  
 complement(join(20992..21064,21164..21274,21355..21500,21554..21715))  
 /gene="ZK678.3"  
 /note="Similarity to Leech cyclin A fragment (TR:G802010)"  
 /codon\_start=1  
 /protein\_id="CAB01905.1"  
 /db\_xref="GI:3881846"  
 /db\_xref="SPTREMBL:Q94412"  
 /translation="MARWLEERTTGGDNFOKETFHLAVSLVDLALPMENIDKMRFO LVGSTSMITKIRGNLPTRNVPDILLMERFLIGKFEVFAKPTPSWLGSCFAKRINLT KMRNDVKLLELSPIDAHLRCPSDAAQSHRINNSTRFISFFMFNKNQFWOCFSFC INI"  
 join(22561..22595,24051..24360,24414..24647)  
 /gene="ZK678.4"  
 join(22561..22595,24051..24360,24414..24647)  
 /gene="ZK678.4"  
 /codon\_start=1  
 /protein\_id="CAB01906.2"  
 /db\_xref="GI:6425514"  
 /db\_xref="SPTREMBL:Q94413"  
 /translation="MGVAVPKNWHLSILOVDKFOITVFQNKYSIILPLSCMFVNIIL HLRFKRETYEMIKKVFSSKHKTITPFAVPLQNTLSKQARRDLIMRQITVSVAVYLS IYELGALLIRVFPTEYAGLPLDLVHKSYFYRYIESIPPTFFVYLETGTSTRMLKRFL KYDGSNFASAAQTVVPVGVAKQQTAVINPHS"  
 23929..24000  
 /gene="ZK678.t3"  
 /note="TGG Pro P-trna"  
 predicted using tRNAscan-SE-1.11  
 preliminary prediction  
 similar to tRNA-Pro"  
 23929..24000

gene  
/gene="ZK678.t3"  
complement(join(24977..25114,25163..25327,25386..25548,25596..25779,25824..25875,26966..27306,27394..27632,28345..28439,28596..28793,29675..29773))  
/gene="ZK678.5"  
/note="vrt-4"  
complement(join(24977..25114,25163..25327,25386..25548,25596..25779,25824..25875,26966..27306,27394..27632,28345..28439,28596..28793,29675..29773))  
/gene="ZK678.5"  
/note="predicted using Genefinder  
contains similarity to Pfam domain: PF01079 (Hint module), Score=391.2, E-value=3.2e-114, N=1  
CDNA EST EMBL:W5796 comes from this gene  
CDNA EST YK105d8.3 comes from this gene  
CDNA EST YK105d8.5 comes from this gene  
CDNA EST YK175d1.3 comes from this gene  
CDNA EST YK175d1.5 comes from this gene  
CDNA EST YK296a11.3 comes from this gene  
CDNA EST YK296a11.5 comes from this gene  
/codon\_start=1  
/protein\_id="CAB01902.2"  
/db\_xref="GI:6425515"  
/db\_xref="SPTREMBL:O94410"  
/translation="MRESLIALVLLSSYKFTYSGCDSTIPYSLEVLSSGPILGC  
ARPCFGHNSGHOLPNAKFRIDQSDGLRDDPLAIHTFDRAADPRVYQAQASCE  
QEPQSLSCNFEDQWVGGIAPVWNASTTKVAICKCTAPLRASIDRGVATVSGGQIVV  
1.7%; Score 19; DB 6; Length 35126;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 ATTATTTGTGTCAGTCAA 40  
|||||  
Db 22708 ATTATTTGTGTCAGTCAA 22726

RESULT 34  
AC004311  
LOCUS  
DEFINITION  
AC004311  
AC004311.1 GI:3769310  
HTG.  
Drosophila melanogaster (Subclones in sac from P1 clone DS07774 (D203)) DNA.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 45105)  
Celniker S.E., George R.A., Galle R.F., Hoskins R.A., Svirkas R.R., Harris N.L., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Chavez C., Chew M., Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim S.H., Lee B., Lomotan M.A., Mak J., Mazda P., Moshrefi A.R., Moshrefi M., Nixon K., Pacleb J.M., Park S., Pfeiffer B., Punch E., Snir E., Twomey B., Wan K.H., Zhang R., Zieran L.L. and Rubin G.M.  
Sequencing of Drosophila chromosome 2R, region 56D11-56E1 Unpublished (1997)  
2 (bases 1 to 45105)  
Celniker S.E., George R.A., Galle R., Svirkas R.R., Hoskins R.A., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Chavez C., Chew M., Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim S.H., Lee B., Lomotan M.A., Mak J., Mazda P., Moshrefi A.R., Moshrefi M., Nixon K., Pacleb J.M., Park S., Pfeiffer B., Punch E., Snir E., Twomey B., Wan K.H., Whitelaw K.R., Yee A., Zhang R., Zieran L.L. and Kimmel B.E.  
Direct Submission  
Submitted (07-MAR-1998) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

Query Match  
Best Local Similarity 1.7%; Score 19; DB 6; Length 35126;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 ATTATTTGTGTCAGTCAA 40  
|||||  
Db 22708 ATTATTTGTGTCAGTCAA 22726

RESULT 34  
AC004311  
LOCUS  
DEFINITION  
AC004311  
AC004311.1 GI:3769310  
HTG.  
Drosophila melanogaster (Subclones in sac from P1 clone DS07774 (D203)) DNA.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 45105)  
Celniker S.E., George R.A., Galle R.F., Hoskins R.A., Svirkas R.R., Harris N.L., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Chavez C., Chew M., Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim S.H., Lee B., Lomotan M.A., Mak J., Mazda P., Moshrefi A.R., Moshrefi M., Nixon K., Pacleb J.M., Park S., Pfeiffer B., Punch E., Snir E., Twomey B., Wan K.H., Zhang R., Zieran L.L. and Rubin G.M.  
Sequencing of Drosophila chromosome 2R, region 56D11-56E1 Unpublished (1997)  
2 (bases 1 to 45105)  
Celniker S.E., George R.A., Galle R., Svirkas R.R., Hoskins R.A., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Chavez C., Chew M., Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim S.H., Lee B., Lomotan M.A., Mak J., Mazda P., Moshrefi A.R., Moshrefi M., Nixon K., Pacleb J.M., Park S., Pfeiffer B., Punch E., Snir E., Twomey B., Wan K.H., Whitelaw K.R., Yee A., Zhang R., Zieran L.L. and Kimmel B.E.  
Direct Submission  
Submitted (07-MAR-1998) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

COMMENT  
On Oct 19, 1998 this sequence version replaced gi:3581706.  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory, MS 64-121  
Berkeley, CA 94720  
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://fruitfly.berkeley.edu/sequence/>) or send email to [bdg@fruitfly.berkeley.edu](mailto:bdg@fruitfly.berkeley.edu).  
Library location: 137-81.

FEATURES  
Source  
1. 45105  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/chromosome="2R"  
/map="56D11-56E1"  
/clone="P1 DS07774 (D203)"  
BASE COUNT 12830 a 9766 c 9450 g 13059 t  
ORIGIN

Query Match  
Best Local Similarity 1.7%; Score 19; DB 4; Length 45105;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 716 AAAGGCACAAAGCCATCCG 734  
|||||  
Db 44238 AAAGGCACAAAGCCATCCG 44256

RESULT 35  
AL451080  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AL451080 56725 bp DNA PRI 02-FEB-2001  
Human DNA sequence from clone RP11-323B13 on chromosome 6, complete sequence.  
AL451080  
AL451080.11 GI:12666307  
HTG.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Clark, S.  
Direct Submission  
Submitted (02-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
On Feb 5, 2001 this sequence version replaced gi:12584477.  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WormPEP; Information on the WormPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> RP11-323B13 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-323B13. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-323B13 is at 1 in this sequence. The true left end of clone RP1-216A11 is at 56626 in this sequence.

## FEATURES

```
Location/Qualifiers
source      1..56725
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="6"
            /clone="RP11-323B13"
            /clone_lib="RPC1-11.2"
            322..608
            /note="AluSx repeat: matches 1..301 of consensus"
repeat_region 642..923
            /note="AluJo repeat: matches 1..295 of consensus"
repeat_region 951..1117
            /note="MIR repeat: matches 1..172 of consensus"
repeat_region 2473..2547
            /note="MIR repeat: matches 185..262 of consensus"
repeat_region 3422..3529
            /note="MER94 repeat: matches 1..111 of consensus"
repeat_region 3544..3643
            /note="L2 repeat: matches 2610..2712 of consensus"
repeat_region 3644..4015
            /note="L2 repeat: matches 1..365 of consensus"
repeat_region 4016..4058
            /note="L2 repeat: matches 2571..2610 of consensus"
repeat_region 4749..5077
            /note="MLT1I repeat: matches 67..410 of consensus"
repeat_region 5149..5456
            /note="AluJb repeat: matches 1..302 of consensus"
repeat_region 6508..6706
            /note="MIR repeat: matches 38..251 of consensus"
repeat_region 7104..7153
            /note="L2 repeat: matches 2045..2095 of consensus"
repeat_region 7203..7714
            /note="L2 repeat: matches 5614..6151 of consensus"
repeat_region 7749..8280
            /note="L2 repeat: matches 2178..2750 of consensus"
repeat_region 8445..8762
            /note="MLT1H repeat: matches 122..420 of consensus"
repeat_region 9173..9648
            /note="MLT1D repeat: matches 1..502 of consensus"
repeat_region 9717..9874
            /note="FRAM repeat: matches 1..157 of consensus"
repeat_region 10235..10680
            /note="LIME repeat: matches 5272..5682 of consensus"
repeat_region 10681..10883
            /note="AluJo repeat: matches 132..309 of consensus"
repeat_region 10884..11193
            /note="AluSx repeat: matches 1..310 of consensus"
repeat_region 11194..11318
            /note="AluJo repeat: matches 1..132 of consensus"
repeat_region 11319..11387
            /note="LIME repeat: matches 5202..5272 of consensus"
repeat_region 11737..12311
            /note="MLT1F repeat: matches 1..541 of consensus"
repeat_region 12553..12881
            /note="MER2 repeat: matches 1..345 of consensus"
repeat_region 13582..13615
            /note="L2 repeat: matches 17..82% conserved"
repeat_region 14281..14579
            /note="AluSx repeat: matches 1..299 of consensus"
repeat_region 14643..14923
            /note="AluJb repeat: matches 1..282 of consensus"
repeat_region 17503..17698
            /note="MLT1D repeat: matches 2..184 of consensus"
repeat_region 17699..17913
            /note="AluSx repeat: matches 59..286 of consensus"
repeat_region 17914..18183
            /note="MLT1D repeat: matches 184..499 of consensus"
repeat_region 18908..19070
            /note="MIR repeat: matches 79..253 of consensus"
            19305..19466
            /note="MIR repeat: matches 48..218 of consensus"
            20101..20402
            /note="AluY repeat: matches 1..307 of consensus"
            21512..21658
            /note="MER69A repeat: matches 5..177 of consensus"
            22024..22134
            /note="L2 repeat: matches 2594..2710 of consensus"
            22647..22839
            /note="L2 repeat: matches 1909..2123 of consensus"
            23563..23857
            /note="AluSx repeat: matches 1..295 of consensus"
            24816..25124
            /note="AluSg repeat: matches 1..304 of consensus"
            25391..25702
            /note="AluJo repeat: matches 1..312 of consensus"
            26974..27220
            /note="MLT1I repeat: matches 121..370 of consensus"
            27349..27386
            /note="L2 repeat: matches 19 copies 2 mer gt 100% conserved"
            27539..28009
            /note="LIME3 repeat: matches 5534..6039 of consensus"
            28166..28365
            /note="MIR repeat: matches 20..231 of consensus"
            28998..29107
            /note="L1MA9 repeat: matches 6155..6270 of consensus"
            30811..32181
            /note="L1MB5 repeat: matches 4089..5490 of consensus"
            32178..33648
            /note="L1PA16 repeat: matches 4635..5619 of consensus"
            33139..33646
            /note="L1PA16 repeat: matches 5643..6156 of consensus"
            33647..34219
            /note="L1MB5 repeat: matches 5471..6065 of consensus"
            34254..34355
            /note="L2 repeat: matches 2172..2274 of consensus"
            35533..35714
            /note="AluY repeat: matches 120..298 of consensus"
            35715..35750
            /note="L2 repeat: matches 18 copies 2 mer ta 94% conserved"
            36662..36730
            /note="MIR repeat: matches 76..145 of consensus"
            37291..37593
            /note="AluSx repeat: matches 1..295 of consensus"
            37601..37912
            /note="AluSx repeat: matches 1..312 of consensus"
            38766..38793
            /note="L2 repeat: matches 14 copies 2 mer tt 100% conserved"
            39638..39857
            /note="MIR repeat: matches 24..262 of consensus"
            40978..41290
            /note="AluSx repeat: matches 1..312 of consensus"
            41291..41655
            /note="L1MB5 repeat: matches 5707..6067 of consensus"
            41656..41962
            /note="AluY repeat: matches 1..307 of consensus"
            41963..42099
            /note="L1MB5 repeat: matches 5570..5707 of consensus"
            42160..42352
            /note="L1MC5 repeat: matches 7744..7925 of consensus"
            42440..42721
            /note="AluSg repeat: matches 16..296 of consensus"
            42738..42917
            /note="L1MC5 repeat: matches 7483..7653 of consensus"
            43173..43268
            /note="MERSA repeat: matches 26..121 of consensus"
            43519..43595
            /note="L1MC3 repeat: matches 7073..7149 of consensus"
            43596..44852
            /note="HSMAR2 repeat: matches 1..1301 of consensus"
            44853..46522
            /note="L1MC3 repeat: matches 5413..7073 of consensus"
```

```

repeat_region 46613..46916
/note="AluX repeat: matches 1..304 of consensus"
repeat_region 47898..48251
/note="THE1B repeat: matches 1..364 of consensus"
repeat_region 48624..48664
/note="L2 repeat: matches 2464..2505 of consensus"
repeat_region 49870..50075
/note="AluSq repeat: matches 1..216 of consensus"
repeat_region 50076..50122
/note="AluSc repeat: matches 243..293 of consensus"
repeat_region 50750..50878
/note="FLAMC repeat: matches 2..129 of consensus"
repeat_region 50889..50936

Query Match 1.7%; Score 19; DB 90; Length 56725;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 GATCACCTTGCTTTTAT 195
|||||
Db 43317 GATCACCTTGCTTTTAT 43335

RESULT 36
AC010019/C
LOCUS AC010019 70534 bp DNA HTG 18-FEB-2000
DEFINITION Drosophila melanogaster clone RPC198-6F6, *** SEQUENCING IN
PROGRESS ***, 37 unordered pieces.
ACCESSION AC010019
VERSION AC010019.3 GI:6996704
KEYWORDS HTG; HTGS_PHASE1.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 70534)
Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
Bodota,B., Bock,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,M., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sugeng,R.,
Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wabba,M.,
Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D., and
Gibbs,R.
Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 70534)
Worley,K.C.
Direct Submission
JOURNAL Submitted (11-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Feb 18, 2000 this sequence version replaced gi:5881475..
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
----- Project name: DBR8
Center project name: DBR8
Center clone name: RPC198-6F6

```

```

----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-primer Bodipy; 21% of reads
Assembly: Dye-terminator Big Dye; 1% of reads
Assembly program: Phrap; version 0.980611
Consensus quality: 35264 bases at least Q40
Consensus quality: 44604 bases at least Q30
Consensus quality: 50877 bases at least Q20
Estimated insert size: 59343; sum-of-contrigs estimation
Quality coverage: 0.4x in Q20 bases; sum-of-contrigs estimation
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1022: contrig of 1022 bp in length
* 1023 1042: gap of unknown length
* 1043 2076: contrig of 1034 bp in length
* 2077 2096: gap of unknown length
* 2097 2907: contrig of 811 bp in length
* 2908 2927: gap of unknown length
* 2928 3801: contrig of 853 bp in length
* 3802 4940: contrig of 1140 bp in length
* 4941 4960: gap of unknown length
* 4961 6246: contrig of 1286 bp in length
* 6247 7251: contrig of 985 bp in length
* 7252 7272: gap of unknown length
* 7273 8216: contrig of 944 bp in length
* 8217 8236: gap of unknown length
* 8237 9531: contrig of 1295 bp in length
* 9532 10409: contrig of 859 bp in length
* 10410 10429: gap of unknown length
* 10430 11088: contrig of 659 bp in length
* 11089 11108: gap of unknown length
* 11109 12041: contrig of 933 bp in length
* 12042 13263: contrig of 1202 bp in length
* 13264 13283: gap of unknown length
* 13284 14108: contrig of 825 bp in length
* 14109 14128: gap of unknown length
* 14129 15376: contrig of 1248 bp in length
* 15377 15396: gap of unknown length
* 15397 16407: contrig of 1011 bp in length
* 16408 16427: gap of unknown length
* 16428 17601: contrig of 1174 bp in length
* 17602 17621: gap of unknown length
* 17622 18555: contrig of 934 bp in length
* 18556 18575: gap of unknown length
* 18576 20269: contrig of 1694 bp in length
* 20270 20289: gap of unknown length
* 20290 21577: contrig of 1288 bp in length
* 21578 21597: gap of unknown length
* 21598 22487: contrig of 890 bp in length
* 22488 22507: gap of unknown length
* 22508 23407: contrig of 900 bp in length
* 23408 23427: gap of unknown length
* 23428 25146: contrig of 1719 bp in length
* 25147 25166: gap of unknown length
* 25167 26802: contrig of 1636 bp in length
* 26803 26822: gap of unknown length
* 26823 28575: contrig of 1753 bp in length
* 28576 28595: gap of unknown length
* 28596 29448: contrig of 853 bp in length
* 29449 29468: gap of unknown length
* 29469 31138: contrig of 1670 bp in length
* 31139 31158: gap of unknown length

```

```

* 31159 32488: contig of 1330 bp in length
* 32489 32508: gap of unknown length
* 32509 34198: contig of 1690 bp in length
* 34199 34218: gap of unknown length
* 34219 36940: contig of 2722 bp in length
* 36941 36960: gap of unknown length
* 36961 38409: contig of 1449 bp in length
* 38410 38429: gap of unknown length
* 38430 40261: contig of 1832 bp in length
* 40262 40281: gap of unknown length
* 40282 41331: contig of 1050 bp in length
* 41332 41351: gap of unknown length
* 41352 44249: contig of 2898 bp in length
* 44250 44269: gap of unknown length
* 44270 47140: contig of 2871 bp in length
* 47141 47160: gap of unknown length
* 47161 48877: contig of 1697 bp in length
* 48878 70534: contig of 21657 bp in length.
FEATURES             Location/Qualifiers
    source            1..70534
    organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone="RPC198-6F6"
BASE COUNT   19607 a 15747 c 15151 g 19302 t      727 others
ORIGIN
Query Match      1.7%; Score 19; DB 61; Length 70534;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1034 AAACATTTTCGGCAGAC 1052
|||||
Db 11911 AAACATTTTCGGCAGAC 11893

RESULT 37
AC016040 70696 bp DNA HTG 10-JAN-2001
LOCUS Homo sapiens chromosome 18 clone RP11-164016 map 18, LOW-PASS
DEFINITION AC016040
ACCESSION AC016040
VERSION AC016040.4 GI:12061546
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 70696)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone RP11-164016
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 70696)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boquelavkiy,L., Boukghalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Hawland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jan 10, 2001 this sequence version replaced gi:9177160.
All repeats were identified using RepeatMasker:

```

```

Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L979
Center clone name: 164_O_16
-----

```

```

* NOTE: This record contains 86 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

```

1
735 834: gap of 100 bp
835 1557: contig of 723 bp in length
1558 1657: gap of 100 bp
1658 2377: contig of 720 bp in length
2378 2477: gap of 100 bp
2478 3193: contig of 716 bp in length
3194 3293: gap of 100 bp
3294 4014: contig of 721 bp in length
4015 4114: gap of 100 bp
4115 4856: contig of 742 bp in length
4857 4956: gap of 100 bp
4957 5695: contig of 739 bp in length
5696 5795: gap of 100 bp
5796 6521: contig of 726 bp in length
6522 6621: gap of 100 bp
6622 7362: contig of 741 bp in length
7363 7462: gap of 100 bp
7463 8182: contig of 720 bp in length
8183 8282: gap of 100 bp
8283 9009: contig of 727 bp in length
9010 9109: gap of 100 bp
9110 9840: contig of 731 bp in length
9841 9940: gap of 100 bp
9941 10671: contig of 731 bp in length
10672 10771: gap of 100 bp
10772 11470: contig of 699 bp in length
11471 11570: gap of 100 bp
11571 12262: contig of 692 bp in length
12263 12362: gap of 100 bp
12363 13071: contig of 709 bp in length
13072 13171: gap of 100 bp
13172 13890: contig of 719 bp in length
13891 13990: gap of 100 bp
13991 14716: contig of 726 bp in length
14717 14816: gap of 100 bp
14817 15544: contig of 728 bp in length
15545 15644: gap of 100 bp
15645 16358: contig of 714 bp in length
16359 16458: gap of 100 bp
16459 17195: contig of 737 bp in length
17196 17295: gap of 100 bp
17296 18028: contig of 733 bp in length
18029 18128: gap of 100 bp
18129 18852: contig of 724 bp in length
18853 18952: gap of 100 bp
18953 19679: contig of 727 bp in length
19680 19779: gap of 100 bp
19780 20538: contig of 759 bp in length
20539 20638: gap of 100 bp
20639 21361: contig of 723 bp in length

```

\* 21362 21461: gap of 100 bp  
\* 21462 22175: contig of 714 bp in length  
\* 22176 22275: gap of 100 bp  
\* 22276 22998: contig of 723 bp in length  
\* 22999 23098: gap of 100 bp  
\* 23099 23823: contig of 725 bp in length  
\* 23824 23923: gap of 100 bp  
\* 23924 24651: contig of 728 bp in length  
\* 24652 24751: gap of 100 bp  
\* 24752 25498: contig of 747 bp in length  
\* 25499 25598: gap of 100 bp  
\* 25599 26328: contig of 730 bp in length  
\* 26329 26428: gap of 100 bp  
\* 26429 27159: contig of 731 bp in length  
\* 27160 27259: gap of 100 bp  
\* 27260 27985: contig of 726 bp in length  
\* 27986 28085: gap of 100 bp  
\* 28086 28810: contig of 725 bp in length  
\* 28811 28910: gap of 100 bp  
\* 28911 29631: contig of 721 bp in length  
\* 29632 29731: gap of 100 bp  
\* 29732 30443: contig of 712 bp in length  
\* 30444 30543: gap of 100 bp  
\* 30544 31242: contig of 699 bp in length  
\* 31243 31342: gap of 100 bp  
\* 31343 32059: contig of 717 bp in length  
\* 32060 32159: gap of 100 bp  
\* 32160 32903: contig of 744 bp in length  
\* 32904 33004: gap of 100 bp  
\* 33004 33732: contig of 729 bp in length  
\* 33733 33832: gap of 100 bp  
\* 33833 34567: contig of 735 bp in length  
\* 34568 34667: gap of 100 bp  
\* 34668 35404: contig of 737 bp in length  
\* 35405 35504: gap of 100 bp  
\* 35505 36246: contig of 742 bp in length  
\* 36247 36346: gap of 100 bp  
\* 36347 37036: contig of 690 bp in length  
\* 37037 37136: gap of 100 bp  
\* 37137 37864: contig of 728 bp in length  
\* 37865 37964: gap of 100 bp  
\* 37965 38658: contig of 694 bp in length  
\* 38659 38758: gap of 100 bp  
\* 38759 39475: contig of 717 bp in length  
\* 39476 39575: gap of 100 bp  
\* 39576 40292: contig of 717 bp in length  
\* 40293 40392: gap of 100 bp  
\* 40393 41093: contig of 701 bp in length  
\* 41094 41193: gap of 100 bp  
\* 41194 41905: contig of 712 bp in length  
\* 41906 42005: gap of 100 bp  
\* 42006 42724: contig of 719 bp in length  
\* 42725 42824: gap of 100 bp  
\* 42825 43562: contig of 738 bp in length  
\* 43563 43662: gap of 100 bp  
\* 43663 44395: contig of 733 bp in length  
\* 44396 44495: gap of 100 bp  
\* 44496 45217: contig of 722 bp in length  
\* 45218 45317: gap of 100 bp  
\* 45318 46033: contig of 716 bp in length  
\* 46034 46133: gap of 100 bp  
\* 46134 46845: contig of 712 bp in length  
\* 46846 46945: gap of 100 bp  
\* 46946 47660: contig of 715 bp in length  
\* 47661 47760: gap of 100 bp  
\* 47761 48475: contig of 715 bp in length  
\* 48476 48575: gap of 100 bp  
\* 48576 49309: contig of 734 bp in length  
\* 49310 49409: gap of 100 bp  
\* 49410 50116: contig of 707 bp in length  
\* 50117 50216: gap of 100 bp  
\* 50217 50940: contig of 724 bp in length  
\* 50941 51040: gap of 100 bp

\* 51041 51749: contig of 709 bp in length  
\* 51750 51849: gap of 100 bp  
\* 51850 52577: contig of 728 bp in length  
\* 52578 52677: gap of 100 bp  
\* 52678 53408: contig of 731 bp in length  
\* 53409 53508: gap of 100 bp  
\* 53509 54213: contig of 711 bp in length  
\* 54220 54319: gap of 100 bp  
\* 54320 55041: contig of 722 bp in length  
\* 55042 55141: gap of 100 bp  
\* 55142 55837: contig of 696 bp in length  
\* 55838 55937: gap of 100 bp  
\* 55938 56675: contig of 738 bp in length  
\* 56676 56775: gap of 100 bp  
\* 56776 57492: contig of 717 bp in length  
\* 57493 57592: gap of 100 bp  
\* 57593 58317: contig of 725 bp in length  
\* 58318 58417: gap of 100 bp  
\* 58418 59148: contig of 731 bp in length  
\* 59149 59248: gap of 100 bp

Query Match 1.7%; Score 19; DB 64; Length 70696;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GATTATTTTGTCTCAGTCA 39

|||||

Db 67139 GATTATTTTGTCTCAGTCA 67157

RESULT 38

AC025735/c

LOCUS

DEFINITION

AC025735

AC025735.4

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AC025735 85472 bp DNA PRI 30-SEP-2000  
Homo sapiens BAC clone RP11-214M24 from Y, complete sequence.  
AC025735

AC025735.4 GI:9454627

HTG.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 85472)

Sulston, J.E. and Waterston, R.

Toward a complete human genome sequence

Genome Res. 8 (11), 1097-1108 (1998)

99063792

2 (bases 1 to 85472)

Kruchowski, S., Maupin, R. and Reitz, L.

The sequence of Homo sapiens BAC clone RP11-214M24

Unpublished

3 (bases 1 to 85472)

Waterston, R.H.

Direct Submission

Submitted (13-MAR-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

4 (bases 1 to 85472)

Waterston, R.H.

Direct Submission

Submitted (26-JUL-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

5 (bases 1 to 85472)

Waterston, R.H.

Direct Submission

Submitted (10-SEP-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

6 (bases 1 to 85472)

Waterston, R.

Direct Submission

Submitted (30-SEP-2000) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

## COMMENT

On Jul 26, 2000 this sequence version replaced gi:7596975.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: sapiens@wustl.wustl.edu  
----- Summary Statistics  
-----  
Center project name: H\_NH0214M24

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skaletsky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

## SOURCE INFORMATION:

The RPCL11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanesi, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

(<http://bacpac.med.buffalo.edu>)

## NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-535113, 200 bp overlap; the clone sequenced to the right is RP11-29003. Actual start of this clone is at base position 32185 of RP11-535113; actual end is at base position 85472 of RP11-214M24.

The clone RP11-214M24 contains a transposon in the vector sequence.

## FEATURES

Source	Location/Qualifiers
	1..85472
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="Y"
	/map="Y"
	/clone="RP11-214M24"
	/clone_lib="RPCL11"
misc_feature	103..298
	/note="similar to EST AW385501 (NID:g6890160)"
misc_feature	146..392
	/note="similar to EST A1025219 (NID:g3240832) ov40g10.x1"
misc_feature	180..298
	/note="similar to EST AW835596 (NID:g7929570)"
repeat_region	424..486
	/rpt_family="L1"
repeat_region	490..861
	/rpt_family="L1"
repeat_region	943..1066
	/rpt_family="L1"
repeat_region	1067..2160
	/rpt_family="ERVK"
repeat_region	2161..3122
	/rpt_family="L1"
repeat_region	3154..3237

repeat_region	/rpt_family="L1"	3278..3528	EST AI383947 (NID:g4196728) te35b12.x1
repeat_region	/rpt_family="L1"	3639..4115	EST AI383947 (NID:g4196728) te35b12.x1
repeat_region	/rpt_family="L1"	4114..5840	EST AW051928 (NID:g5914287) wx24c04.x1
repeat_region	/rpt_family="L1"	5880..6180	EST D20390 (NID:g501486)"
misc_feature	/rpt_family="Alu"	6191..6422	EST AI701149 (NID:g4989049) we10e10.x1
repeat_region	/note="similar to	6452..6768	EST AA769099 (NID:g2820337) oa74f12.s1
misc_feature	/rpt_family="Alu"	6783..6990	EST AI383947 (NID:g4196728) te35b12.x1
misc_feature	/note="similar to	6791..7289	EST AW051928 (NID:g5914287) wx24c04.x1
misc_feature	/note="similar to	6819..6990	EST D20390 (NID:g501486)"
misc_feature	/note="similar to	6863..7454	EST AI701149 (NID:g4989049) we10e10.x1
misc_feature	/note="similar to	7029..7585	EST AA769099 (NID:g2820337) oa74f12.s1
misc_feature	/note="similar to	7076..7590	EST AI383947 (NID:g4196728) te35b12.x1
misc_feature	/note="similar to	7160..7585	EST D20390 (NID:g501486)"
misc_feature	/note="similar to	7169..7438	EST AI701149 (NID:g4989049) we10e10.x1
misc_feature	/note="similar to	7200..7407	EST AA769099 (NID:g2820337) oa74f12.s1
misc_feature	/note="similar to	7265..7756	EST AI383947 (NID:g4196728) te35b12.x1
misc_feature	/note="similar to	7335..7772	EST AW051928 (NID:g5914287) wx24c04.x1
misc_feature	/note="similar to	7692..7897	EST D20390 (NID:g501486)"
repeat_region	/note="similar to	8401..8694	EST AI701149 (NID:g4989049) we10e10.x1
misc_feature	/rpt_family="Alu"	8811..9479	EST AA769099 (NID:g2820337) oa74f12.s1
misc_feature	/note="similar to	9236..9643	EST AI383947 (NID:g4196728) te35b12.x1
misc_feature	/note="similar to	9252..9643	EST D20390 (NID:g501486)"
misc_feature	/note="similar to	9288..9643	EST AI701149 (NID:g4989049) we10e10.x1
repeat_region	/note="similar to	11767..11922	EST AA769099 (NID:g2820337) oa74f12.s1
repeat_region	/rpt_family="ERV1"	11986..12080	EST AI383947 (NID:g4196728) te35b12.x1
repeat_region	/rpt_family="Alu"	12084..12403	EST AW051928 (NID:g5914287) wx24c04.x1
repeat_region	/rpt_family="ERV1"	13258..13562	EST D20390 (NID:g501486)"
repeat_region	/rpt_family="Alu"	13607..13910	EST AI701149 (NID:g4989049) we10e10.x1
repeat_region	/rpt_family="Alu"	14815..14990	EST AA769099 (NID:g2820337) oa74f12.s1
repeat_region	/rpt_family="L1"	15171..15455	EST AI383947 (NID:g4196728) te35b12.x1
repeat_region	/rpt_family="Alu"	15467..15680	EST D20390 (NID:g501486)"
repeat_region	/rpt_family="L1"	15706..16127	EST AI701149 (NID:g4989049) we10e10.x1
repeat_region	/rpt_family="ERV1"	16128..17624	EST AA769099 (NID:g2820337) oa74f12.s1
repeat_region	/rpt_family="ERV1"	17659..17774	EST AI383947 (NID:g4196728) te35b12.x1
repeat_region	/rpt_family="L1"	17779..18803	EST AW051928 (NID:g5914287) wx24c04.x1
repeat_region	/rpt_family="L1"	18826..18928	EST D20390 (NID:g501486)"
repeat_region	/rpt_family="BSR/beta"		

```

repeat_region 19533..19690
/rpt_family="BSR/beta"
repeat_region 19754..19838
/rpt_family="BSR/beta"
repeat_region 19897..20175
/rpt_family="BSR/beta"
misc_feature 20620..21023
/Note="similar to EST AI026871 (NID:g3246359) ov92a08.xl"
repeat_region 21361..21477
/rpt_family="BSR/beta"
repeat_region 21513..21582
/rpt_family="BSR/beta"
repeat_region 22587..22718

Query Match 1.7%; Score 19; DB 88; Length 85472;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 871 GTCACAGAAGTCCTGGCTT 889
|||||
Db 22395 GTCACAGAAGTCCTGGCTT 22377
|||||

RESULT 39
AC016550/c 88574 bp DNA PRI 19-DEC-2000
LOCUS Homo sapiens chromosome 5 clone CTC-261N6, complete sequence.
AC016550
AC016550.5 GI:11890782
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 88574)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS
Direct Submission
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 88574)
DOE Joint Genome Institute.
AUTHORS
Direct Submission
JOURNAL
Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 88574)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS
Direct Submission
JOURNAL
Submitted (19-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Dec 19, 2000 this sequence version replaced gi:7710243.
COMMENT
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.7% of Sequence;
Estimated Total Number of Errors is 0.2.
FEATURES
Location/Qualifiers
source
1..88574
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-261N6"
BASE COUNT 30699 a 15219 c 14549 g 28107 t
ORIGIN

Query Match 1.7%; Score 19; DB 87; Length 88574;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 479 TTGTTGTGGGGCTGTC 497
|||||
Db 81775 TTGTTGTGGGGCTGTC 81757
|||||

```

```

RESULT 40
AC005148
LOCUS Drosophila melanogaster DNA sequence (Pls DS05759 (D204) and
DEFINITION DS09033 (D211)), complete sequence.
AC005148 AC004312 AC003593 AC003594 AC003852 AC003595 AC003596
AC003853 AC003597 AC003598 AC004317 AC003601
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (Subclones in tet from P1 clones DS05759
(D204) and DS09033 (D211)) DNA.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 95262)
REFERENCE
AUTHORS
Celniker, S.E., Aghavani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Chew, M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A.,
Humast, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B.,
Lomotan, M.A., Mazda, P., Mok, M.S., Moshrefi, A.R., Moshrefi, M.,
Nixon, K., Pacleib, J.M., Park, S., Pfeiffer, B., Punch, D., Snir, E.,
Twomey, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.L.
and Kimmel, B.
TITLE
Sequencing of Drosophila chromosome 2L, region 56D11-56E2.
JOURNAL
Unpublished (1997)
REFERENCE
2 (bases 1 to 95262)
Celniker, S.E., Aghavani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Chew, M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A.,
Humast, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B.,
Lomotan, M.A., Mazda, P., Mok, M.S., Moshrefi, A.R., Moshrefi, M.,
Nixon, K., Pacleib, J.M., Park, S., Pfeiffer, B., Punch, D., Snir, E.,
Twomey, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.L.
and Kimmel, B.
TITLE
Direct Submission
JOURNAL
Submitted (20-JUN-1998) Berkeley Drosophila Genome Project, MS
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US
COMMENT
Sequence submitted by:
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site
(http://fruitfly.berkeley.edu/sequence-archive.html) or send email
to drosophila@hgsc.lbl.gov.
Library locations: 138,60, 9,95.
FEATURES
Location/Qualifiers
source
1..95262
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2R"
/map="56D11-56E2"
/clone="Pls DS05759 (D204) and DS09033 (D211)"
/Note="DS05759 (D204) extends from bp 1 to bp 81035 and
DS09033 (D211) extends from bp 80404 to bp 95264"
BASE COUNT 28107 a 18878 c 19928 g 28349 t
ORIGIN

Query Match 1.7%; Score 19; DB 4; Length 95262;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 716 AAAGGCACAAAGCCATCCG 734
|||||
Db 1610 AAAGGCACAAAGCCATCCG 1628
|||||

RESULT 41
AF228729
LOCUS AF228729 106356 bp DNA HTG 01-JUN-2000

```

Homo sapiens chromosome 8 clone RPL1-16A9, WORKING DRAFT SEQUENCE,  
10 unordered pieces.

DEFINITION AF228729  
ACCESSION AF228729.2 GI:8151964  
VERSION  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Schudy A., Blechschmidt K., Menzel U., Polley A., Reichwald K.,  
Rump A., Schilhabel M.B., Taudien S., Wen G. and Rosenthal A.  
TITLE Chromosome 8 genomic sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 106356)  
AUTHORS Schudy A., Schilhabel M., Baumgart C., Menzel U., Weber J.,  
Schattevoy R. and Rosenthal A.  
TITLE Direct Submission  
JOURNAL Submitted (25-JAN-2000) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
REFERENCE 3 (bases 1 to 106356)  
AUTHORS Genome Sequencing Center Jena.  
JOURNAL Direct Submission  
TITLE Submitted (24-MAY-2000) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
JOURNAL On Jun 1, 2000 this sequence version replaced gi:5960206.  
COMMENT \* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

1 3655: contig of 3655 bp in length  
\* 3656 3755: gap of unknown length  
\* 3756 30944: contig of 27189 bp in length  
\* 30945 31044: gap of unknown length  
\* 31045 53141: contig of 22097 bp in length  
\* 53142 53241: gap of unknown length  
\* 53242 66755: contig of 13514 bp in length  
\* 66756 66855: gap of unknown length  
\* 66856 75196: contig of 8341 bp in length  
\* 75197 75296: gap of unknown length  
\* 75297 78918: contig of 3622 bp in length  
\* 78919 79018: gap of unknown length  
\* 79019 90860: contig of 11842 bp in length  
\* 90861 90961: gap of unknown length  
\* 90962 97433: contig of 6472 bp in length  
\* 97433 97532: gap of unknown length  
\* 97533 98555: contig of 1023 bp in length  
\* 98556 98555: gap of unknown length  
\* 98556 106356: contig of 7701 bp in length.

FEATURES  
source Location/Qualifiers  
1..106356  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/clone="RPL1-16A9"  
misc\_feature 1..3655  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"  
misc\_feature 97756..106356  
/note="assembly\_fragment  
clone\_end:T7  
vector\_side:right"  
BASE COUNT 31212 a 21597 c 21675 g 30794 t 1078 others  
ORIGIN

Query Match 1.7%; Score 19; DB 78; Length 106356;  
Best Local Similarity 100.0%; Pred. No. 47;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 811 AATTGGGTAATAATCAACC 829  
|||||  
DB 22499 AATTGGGTAATAATCAACC 22517

RESULT 42  
HS347H13  
LOCUS  
DEFINITION Human DNA sequence from clone RP3-347H13 on chromosome 22 Contains  
the 3' end of the ACO2 gene for mitochondrial aconitase 2, the gene  
for a novel protein similar to yeast DNA-directed RNA polymerase  
III 25 kD polypeptide, the gene for the human ortholog of rat  
Pippin, the PMM1 gene for phosphomannomutase 1 (PMMH-22, yeast  
SEC53 ortholog), and two novel genes. Contains ESTs, STSS, GSSs and  
two putative CpG islands, complete sequence.  
AL023553  
ACCESSION AL023553.5 GI:4490860  
VERSION  
KEYWORDS HTG; ACO2; aconitase; CpG island; phosphomannomutase; Pippin; PMM1;  
PMMH-22; RNA Polymerase III; SEC53.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 107818)  
TITLE Direct Submission  
JOURNAL Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Mar 23, 1999 this sequence version replaced gi:4490464.  
COMMENT This sequence has been finished according to sequence map criteria  
as follows. An attempt is made to resolve all sequencing problems,  
such as compressions and repeats, but not necessarily within known  
annotated human repeat sequence elements (e.g. Alu). Where the  
sequence is ambiguous, there is an annotation using the 'unsure'  
feature key.  
This sequence was generated from part of bacterial clone contigs of  
human chromosome 22, constructed by the Sanger Centre Chromosome 22  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr22  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP3-347H13 is  
from the library RPCI-3 constructed at the Roswell Park Cancer  
Institute by the group of Pieter de Jong. For further details see  
http://bacpac.med.buffalo.edu/  
VECTOR: PCYPAC2  
IMPORTANT: This sequence is not the entire insert of clone  
RP3-347H13. It may be shorter because we only sequence overlapping  
sections once, or longer because we arrange for a small overlap  
between neighbouring submissions.  
The true left end of clone RP3-347H13 is at 1 in this sequence.  
The true left end of clone CTA-216E10 is at 107715 in this sequence.  
The true right end of clone CTA-223H9 is at 13652 in this sequence.  
The start of this sequence overlaps with sequence Z83840 The end of  
this sequence overlaps with sequence AL008582.

FEATURES  
source Location/Qualifiers  
1..107818  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="RP3-347H13"  
/clone\_lib="RPCI-3"

repeat_region	1..237	/note="AluSg repeat: matches 1..235 of consensus"	gene	/product="dj347hl3.1 (aconitase 2, mitochondrial
repeat_region	330..473	/note="AluJo repeat: matches 1..149 of consensus"		{Aconitate Hydratase, EC 4.2.1.3, Citrate Hydrolyase})"
repeat_region	474..768	/note="AluSx repeat: matches 5..299 of consensus"		Join(2595..2731,10660..10918,14745..14837,18246..18404,
repeat_region	771..1035	/note="AluSc repeat: matches 39..303 of consensus"		18636..18786,20396..20500,21350..21441,23040..23145,
repeat_region	1036..1194	/note="AluJo repeat: matches 149..296 of consensus"		25699..25856,26074..26147,26699..26810,27715..27837,
repeat_region	1222..1680	/note="MLRIF repeat: matches 111..538 of consensus"		28062..28217,29131..29322,30157..30289,30770..30891,
repeat_region	1681..1976	/note="AluSx repeat: matches 1..298 of consensus"		31348..31858)
repeat_region	1977..2012	/note="MLRIF repeat: matches 78..111 of consensus"		/gene="ACO2"
repeat_region	2072..2181	/note="L2 repeat: matches 2614..2733 of consensus"		Join(<2595..2731,10660..10918,14745..14837,18246..18404,
repeat_region	2427..2471	/note="MIR repeat: matches 117..162 of consensus"		18636..18786,20396..20500,21350..21441,23040..23145,
mRNA		Join(<2595..2731,10660..10918,14745..14837,18246..18404,		25699..25856,26074..26147,26699..26810,27715..27837,
		28062..28217,29131..29322,30157..30289,30770..30891,		31348..31858)
		/gene="ACO2"		/translating="KALGVQYHVASVLCORAKVAMSHFEPNEYIHYDLLEKNINIVR
		/note="match: cDNAs: Em:J05224 Em:AF186471 Em:U800404		KRLNPLTSEKIVYCHLDDPASOETRGKSYLRLRDRVAMODATAQWAMLOFISGG
		Em:249931 Em:AF086790 Em:278147		LSKVAVPSTIHCDHLIEAQVGGKDLRRKINDQYVNFATAGAKYGVGFWKPGSGI
		match: ESTs: Em:AA840484 Em:W14196 Em:AA274874 Em:AA597053		GVKLGTSLSGSSPKDKVLKVGITGTAIVYHGVGVDISICTGMATCNMGA
		Em:AA173509 Em:N36454 Em:AA930989 Em:A1037103 Em:A1346207		ETGATSVFVYNNRMKKYKSGTREDIANLAFEDKHLVPDPCGHDQIETINSELK
		Em:AA273346 Em:A1506948 Em:AA560856		PHNGKFPDLHVAEKGKGLDRLVGLIGSTNCCYEDMGSAVAKOAL
		Em:AA797850 Em:A1459540 Em:AA791448 Em:A160630 Em:H05990		AHLGKQSOFTTTPGSEQIRATIRDEGVAQILRLDGLGIVLANACSGICGOWDRKDLK
		Em:AA756194 Em:A2460284 Em:AA9530230 Em:A1118987		CEKNIVTSYNNFTGRNDANETHAFVTSPEIVTALATAGTLKFNPEYDYLGTGCK
		Em:AA895480 Em:H14705 Em:AA791541 Em:AA359441 Em:H37881		LEDKAPADELPKGFDPQDTYQHPKDDSGQHVDSPTSROLLEPPFDKWDGKD
		Em:H16897 Em:AA530429 Em:AA724140 Em:AA474367 Em:AA2339872		KRLQALIKVKGKCTDHISAAGPMLKFRGLDNTSNLLIGAINENKANSVRNAV
		Em:AA670548 Em:AA815588 Em:T77344 Em:A1304443 Em:AA724144		TEFGFVPDPTARYKKHGVVIGDENTGESSREHAALPRHLGRAITIKTSFARI
		Em:A624931 Em:C03874 Em:AA591946 Em:A153209 Em:R69407		HETNLKQGLPLTTPADPADYNNKIHPVDKLTIQGLKFTPGKPLKCIKIHNPQTETI
		Em:T08348 Em:AA644832 Em:N36629 Em:AA833336 Em:R87982		LLNHTFNETQIEWFRAGSALNMKELQQ"
		Em:R77967 Em:AA776120 Em:AA410123 Em:R13530 Em:AA791264		3159..3208
		Em:AA66067 Em:AA530353 Em:R17339 Em:298458 Em:W30158		/note="25 copies 2 mer tt 74 conserved"
		Em:R60747 Em:R88256 Em:AA834335 Em:AA073622 Em:A1225733		3269..3553
		Em:AA103640 Em:AA275929 Em:A1060658 Em:AA105776		/note="AluSg repeat: matches 1..283 of consensus"
		Em:AA775823 Em:AA530778 Em:AA351027 Em:A1077293		3569..3872
		Em:AA259860 Em:F07688 Em:R88847 Em:F01149 Em:AA560345		/note="AluYb8 repeat: matches 1..314 of consensus"
		Em:T31100 Em:AA538092 Em:AA007867 Em:R56675 Em:AA637285		3880..4170
		Em:AA334769 Em:H61013 Em:AA250646 Em:W40671 Em:N25997		/note="AluJo repeat: matches 1..293 of consensus"
		Em:AA337850 Em:AA507744 Em:N25740 Em:AA312812 Em:A1214043		4175..4231
		Em:AA259859 Em:AA384173 Em:AA311248 Em:AA048286 Em:T32036		/note="MIR repeat: matches 83..143 of consensus"
		Em:AA117755 Em:A1119350 Em:AA449580 Em:AA563956 Em:R74670		4810..5006
		Em:N25765 Em:A1041836 Em:T33408 Em:HI5984 Em:F06809		/note="MIR repeat: matches 20..219 of consensus"
		Em:AA363904 Em:F01087 Em:A1347998 Em:AA511772 Em:R56700		5007..5107
		Em:A1439268 Em:T31967 Em:AA914534 Em:T19635 Em:AA084544		/note="MER5B repeat: matches 75..176 of consensus"
		Em:A1159781 Em:AA008868 Em:T31555 Em:A1079766 Em:AA160535		5108..5322
		Em:AA315324 Em:A114955 Em:AA994289 Em:AA087364 Em:H41420		/note="MER5A repeat: matches 1..189 of consensus"
		Em:C05156 Em:AA034740 Em:R7156 Em:A1439314 Em:AA477990		5466..5778
		Em:A1373134 Em:MG2227 Em:T19636 Em:T39113 Em:AA778441		/note="AluSx repeat: matches 1..312 of consensus"
		Em:A1380230 Em:AA832135 Em:AA10706733 Em:AA062620 Em:T30933		5779..5913
		Em:F08310 Em:A1302579 Em:A1085953 Em:AA064974 Em:T34603		/note="AluJo/FLAM repeat: matches 2..132 of consensus"
		Em:H78528 Em:N33768 Em:H97846 Em:A1235320 Em:AA580271		6190..6497
		Em:N86127 Em:221666 Em:W08849 Em:H51364 Em:H84459		/note="AluSg repeat: matches 1..304 of consensus"
		Em:HI19923 Em:AA442333 Em:A1083635 Em:AA730127 Em:AA572160		6525..6749
		Em:AA477380 Em:H46787 Em:AA843583 Em:AA399651 Em:T94732		/note="MER20 repeat: matches 2..218 of consensus"
		Em:T60001 Em:A594044 Em:H23758 Em:A1148661 Em:241314		6801..7114
		Em:F19517 Em:A1201643 Em:AA633505 Em:AA852447 Em:R33460		/note="AluX repeat: matches 1..298 of consensus"
		Em:H484847 Em:W00583 Em:AA648227 Em:AA130198 Em:A1475036		
		Em:A1283817 Em:A1475773 Em:AA324903 Em:239392 Em:AA333992		
		Em:H55171 Em:A1032594 Em:H14600 Em:H93762 Em:AA134095		
		Em:N89324 Em:A1081081 Em:R56787 Em:AA661740 Em:N89374		
		Em:R61708 Em:R52552 Em:AA648642 Em:H23571 Em:A112740		
		Em:A1379801 Em:F21578 Em:A1345885 Em:A1377254 Em:A1236736		
		Em:AA449323 Em:H45170"		
		/evidence-not_experimental		
			Query Match	1.7%; Score 19; DB 91; Length 107818;
			Best Local Similarity	100.0%; Pred. No. 47;
			Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
			QY	143 GTGTTCTTGGCTCCGCGG 161
			DB	40737 GGTGTTCTTGGCTCCGCGG 40755

## RESULT 43

AC005846/c  
 AC005846 111904 bp DNA PRI 27-JUN-2000  
 Homo sapiens 12p13.3 PAC RPCI3-467F14 (Roswell Park Cancer  
 Institute Human PAC Library) complete sequence. Contains Human von  
 Willebrand factor gene exons 4 through 34.  
 AC005846 AC006523  
 AC005846.1 GI:4092459  
 HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 111904)  
 Muzny, D., Arenson, A.D., Bouck, J., Bunac, C., Chen, Z., Ding, Y.,  
 Dugan, S., Durbin, J., Forcum, J., Garcia, C., Gorrell, J.H.,  
 Gorrell, L.B., Hernandez, J., Issar, A., Jackson, L., Kneitz, S.,  
 Kondrjewski, N., Lau, S., Leal, B., Lee, E., Lichtarge, O., Liu, W.,  
 Logan, O., Lu, J., Marondel, I., Martinez, C., Merscher, S., Miller, A.,  
 Montgomery, K., Oswal, G., Pampell, L.R., Parish, B.J., Perez, L.,  
 Rashid, N.D., Rives, C., Scherer, S.E., Shen, H., Shim, C., Simon, M.,  
 Vo, Q., Williamson, A., Worley, K.C., Xhang, A.M., Yang, R., Yu, W.,  
 Zhou, X., Kucherlapati, R., Nelson, D. and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 111904)  
 Worley, K.C.  
 Direct Submission  
 Submitted (22-OCT-1998) Molecular and Human Genetics, Baylor  
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 111904)  
 Worley, K.C.  
 Direct Submission  
 Submitted (01-JAN-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 4 (bases 1 to 111904)  
 Worley, K.C.  
 Direct Submission  
 Submitted (08-JAN-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 5 (bases 1 to 111904)  
 Worley, K.C.  
 Direct Submission  
 Submitted (27-JUN-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jun 27, 2000 this sequence version replaced gi:4803879  
 gi:4033630.  
 INFORMATION: <http://gc.bcm.tmc.edu:8088/home.html> or email  
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the  
 entire insert of this clone. Overlapping regions of clones are only  
 sequenced and submitted once, so the sequence for the remainder of  
 the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the  
 features listing.

## ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches  
 of a local database that includes entries from dbSTS, GDB, and  
 local mapping efforts.  
 Repeats are identified using RepeatMasker (A. Smit and P. Green,  
 unpublished.) for Human and Mouse sequences.  
 Genes and Region of sequence similarity are identified by BLAST  
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
 EST and cDNA sequences. Genes demonstrate at least two exons  
 flanked by consensus splice sites that maintained sequence  
 continuity across the splice junctions. Sequences that are not  
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
 standard of double strand coverage with a minimum of 2 clones and 2  
 reads with no ambiguities or 2 chemistries with a minimum of 2  
 clones and 3 reads with no ambiguities. If the sequence quality for  
 a region does not meet this standard, it will be indicated in the  
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
 standards - estimated error rate less than 1 per 10,000 bases.  
 Reports of lowest quality individual bases and measures of base  
 quality are listed below. Description of the metrics can be found  
 at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT-----

----- Summary Statistics -----  
 Contig length: 94462  
 Phrap values in estimate: 93329  
 Average error rate (BCM-Phrap estimate): 0.000923054  
 Fraction of Phrap values less than 40 : 0.045195  
 Number of consensus changing edits: 3  
 Number of N's in consensus : 0

----- Consensus changing edits -----  
 Position Original+Context Edited+Context  
 19579 agcgcttaa(n)cactgcgttc agcgcttaa(t)cactgcgttc  
 29991 tttagtaga(a)atggggttc tttagtaga(g)atggggttc  
 89256 gcatgggtct(n)gcttcgccca gcatgggtct(g)gcttcgccca

## ----- Distribution of Quality &lt; 40 Bases -----

#	bases	5	10	15	20	25	30	35	40
1000									
900									*
800									*
700									*
600									*
500									*
400									*
300				*	*	*	*	*	*
200			*	*	*	*	*	*	*
100		*	*	*	*	*	*	*	*
0		*	*	*	*	*	*	*	*

Phrap Value Range

Version: 1.01 qxfo.  
 Location/Qualifiers  
 1. 111904  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12p13.3"  
 /clone="RPCI3-467F14"  
 43. 319  
 /rpt\_family="AluX"  
 complement(335..441)  
 /rpt\_family="MER31A"  
 complement(783..1081)  
 /rpt\_family="AluX"  
 complement(1109..1249)  
 /rpt\_family="L2"  
 complement(1250..1412)  
 /rpt\_family="MER5A"  
 1962..2258  
 /rpt\_family="AluX"  
 complement(2816..3068)  
 /rpt\_family="AluSp"

```
repeat_region 4761..4911
repeat_region /rpt_family="MER20"
repeat_region 5307..5604
repeat_region /rpt_family="AluJo"
repeat_region 5990..6043
repeat_region /rpt_family="MIR"
repeat_region /complement(6932..7020)
repeat_region /rpt_family="L1MC4"
repeat_region 7053..7347
repeat_region /rpt_family="AluY"
repeat_region /complement(7736..7841)
repeat_region /rpt_family="(CA)n"
repeat_region /complement(8668..8976)
repeat_region /rpt_family="AluY"
repeat_region 9030..9259
repeat_region /rpt_family="AluDb"
repeat_region /complement(9779..9779,10063..10271,125061..25185,45094..45310,
gene 46904..47026,48203..48314,49302..49348,55372..55508,
56261..56399,57591..57692,62601..62796,63577..63788,
67862..68102,73828..73922,76194..76354,84154..84257,
85819..85957,89067..89201,91157..91303,94599..94739,
94952..95065,96858..97014,97747..97905,98610..98745,
100902..102280,103775..103891,103989..104129,
104413..104556,107000..107164,108515..108558,
108851..109028)
/rpt_family="Human von Willebrand factor gene, M25865."
complement(9835..9910)
repeat_region /rpt_family="MIR"
repeat_region 10696..10991
repeat_region /rpt_family="AluY"
repeat_region /complement(11228..11294)
repeat_region /rpt_family="L1M4"
repeat_region /complement(11428..11542)
repeat_region /rpt_family="FLAM_A"
repeat_region /complement(11646..11945)
repeat_region /rpt_family="AluY"
repeat_region /complement(12278..12316)
repeat_region /rpt_family="(CA)n"
repeat_region /complement(12332..12633)
repeat_region /rpt_family="AluSx"
misc_feature 12677..12817
repeat_region /function="Low coverage"
repeat_region 12888..12974
repeat_region /rpt_family="AluY"
repeat_region /complement(13020..13203)
repeat_region /rpt_family="L1M4"
repeat_region /complement(13222..13689)
repeat_region /rpt_family="L1ME1"
repeat_region /complement(13750..14053)
repeat_region /rpt_family="AluSg"
repeat_region /complement(14054..14412)
repeat_region /rpt_family="L1MB3"
repeat_region /complement(14440..14585)
repeat_region /rpt_family="L1MB3"
repeat_region /complement(14586..14889)
repeat_region /rpt_family="AluSc"

Query Match 1.7% Score 19; DB 86; Length 111904;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 TGCCACGTGCAAGTTGCTA 336
|||||
Db 10519 TGCCACGTGCAAGTTGCTA 10501

RESULT 44
AC026444/c AC026444 126663 bp DNA HTG 18-JUL-2000
LOCUS Homo sapiens chromosome 5 clone CTD-2299E8, WORKING DRAFT SEQUENCE,
DEFINITION 19 ordered pieces.
ACCESSION AC026444
VERSION AC026444.3 GI:9256515
```

```
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

```
HTG: HTGS_PHASE2; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 126663)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 126663)
DOE Joint Genome Institute.
Direct Submission
Submitted (22-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:7711916.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 742469, H459
Center clone name: CITB-HI_2299E8
-----
Summary Statistics
Consensus quality: 114752 bases at least Q40
Consensus quality: 122704 bases at least Q30
Consensus quality: 124065 bases at least Q20
Estimated insert size: 121000; pulse field gel estimation
Estimated insert size: 125813; sum-of-contigs estimation
Quality coverage: 4.44 in Q20 bases; pulse field gel estimation
Quality coverage: 4.27 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 7518: contig of 7518 bp in length
7519 7618: gap of unknown length
7619 10622: contig of 3004 bp in length
10623 10722: gap of unknown length
10723 13537: contig of 2815 bp in length
13538 13637: gap of unknown length
13638 18289: contig of 4652 bp in length
18290 18389: gap of unknown length
18390 20907: contig of 2518 bp in length
20908 21007: gap of unknown length
21008 23015: contig of 2008 bp in length
23016 23115: gap of unknown length
23116 27055: contig of 3940 bp in length
27056 27155: gap of unknown length
27156 30501: contig of 3346 bp in length
30502 32694: contig of 2093 bp in length
32695 36294: gap of unknown length
36295 36621: contig of 3827 bp in length
36622 36721: gap of unknown length
36722 40262: contig of 3541 bp in length
40263 40362: gap of unknown length
40363 44764: contig of 4402 bp in length
44765 44864: gap of unknown length
44865 47712: contig of 2848 bp in length
47713 47812: gap of unknown length
47813 82440: contig of 34628 bp in length
82441 92580: contig of 10040 bp in length
92581 92680: gap of unknown length
92681 96611: contig of 3931 bp in length
96612 96711: gap of unknown length
```

```
HTG: HTGS_PHASE2; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 126663)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 126663)
DOE Joint Genome Institute.
Direct Submission
Submitted (22-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:7711916.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 742469, H459
Center clone name: CITB-HI_2299E8
-----
Summary Statistics
Consensus quality: 114752 bases at least Q40
Consensus quality: 122704 bases at least Q30
Consensus quality: 124065 bases at least Q20
Estimated insert size: 121000; pulse field gel estimation
Estimated insert size: 125813; sum-of-contigs estimation
Quality coverage: 4.44 in Q20 bases; pulse field gel estimation
Quality coverage: 4.27 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 7518: contig of 7518 bp in length
7519 7618: gap of unknown length
7619 10622: contig of 3004 bp in length
10623 10722: gap of unknown length
10723 13537: contig of 2815 bp in length
13538 13637: gap of unknown length
13638 18289: contig of 4652 bp in length
18290 18389: gap of unknown length
18390 20907: contig of 2518 bp in length
20908 21007: gap of unknown length
21008 23015: contig of 2008 bp in length
23016 23115: gap of unknown length
23116 27055: contig of 3940 bp in length
27056 27155: gap of unknown length
27156 30501: contig of 3346 bp in length
30502 32694: contig of 2093 bp in length
32695 36294: gap of unknown length
36295 36621: contig of 3827 bp in length
36622 36721: gap of unknown length
36722 40262: contig of 3541 bp in length
40263 40362: gap of unknown length
40363 44764: contig of 4402 bp in length
44765 44864: gap of unknown length
44865 47712: contig of 2848 bp in length
47713 47812: gap of unknown length
47813 82440: contig of 34628 bp in length
82441 92580: contig of 10040 bp in length
92581 92680: gap of unknown length
92681 96611: contig of 3931 bp in length
96612 96711: gap of unknown length
```

\* 96712 111639: contig of 14928 bp in length  
 \* 111640 111739: gap of unknown length  
 \* 111740 125113: contig of 13373 bp in length  
 \* 125113 125213: gap of unknown length  
 \* 125213 126663: contig of 1451 bp in length.

## FEATURES

## source

1. 126663  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="CWD-2299E8"  
 /clone\_lib="Caltech human BAC library D"  
 BASE COUNT 41814 a 22538 c 22035 g 38469 t 1807 others  
 ORIGIN

Query Match 1.78; Score 19; DB 70; Length 126663;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 479 TTGTTGTGGGCTGC 497

Db 8907 TTGTTGTGGGCTGC 8889

## RESULT 45

AC012363 128133 bp DNA HTG 17-AUG-2000  
 LOCUS Homo sapiens chromosome 2 clone RP11-438O12, WORKING DRAFT  
 DEFINITION

SEQUENCE, 36 unordered pieces.

## ACCESSION

AC012363

Submitted (25-OCT-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Aug 17, 2000 this sequence version replaced gi:8440000.

## REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

2 (bases 1 to 128133)

Waterston,R.H.

Direct Submission

Submitted (25-OCT-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Aug 17, 2000 this sequence version replaced gi:8440000.

## COMMENT

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 ----- Project Information -----  
 Center project name: H\_NH0438012  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 67%  
 Chemistry: Dye-terminator; 33%  
 Chemistry: Dye-terminator Big Dye; 33% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 182549 bases at least Q40  
 Consensus quality: 190814 bases at least Q30  
 Consensus quality: 193372 bases at least Q20  
 Insert size: 218000; agarose-fp  
 Quality coverage: 3.64 in Q20 bases; agarose-fp  
 Quality coverage: 3.84 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 36 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 770: contig of 770 bp in length  
 771: gap of unknown length  
 871: contig of 1245 bp in length  
 2115: gap of unknown length  
 2215: gap of 831 bp in length  
 3046: gap of unknown length  
 3146: gap of 878 bp in length  
 4024: contig of 878 bp in length  
 4124: gap of unknown length  
 5010: contig of 886 bp in length  
 5110: gap of unknown length  
 5111: contig of 1337 bp in length  
 6447: gap of unknown length  
 6547: contig of 1366 bp in length  
 7913: gap of unknown length  
 8013: contig of 1646 bp in length  
 9659: gap of unknown length  
 9759: gap of unknown length  
 11026: contig of 1267 bp in length  
 11126: gap of unknown length  
 12216: contig of 1090 bp in length  
 12316: gap of unknown length  
 12509: contig of 193 bp in length  
 12609: gap of unknown length  
 14475: contig of 1866 bp in length  
 14575: gap of unknown length  
 16865: contig of 2290 bp in length  
 16965: gap of unknown length  
 19653: contig of 2688 bp in length  
 19753: gap of unknown length  
 21718: contig of 1965 bp in length  
 21818: gap of unknown length  
 24048: contig of 2230 bp in length  
 24148: gap of unknown length  
 27164: contig of 3016 bp in length  
 27264: gap of unknown length  
 30695: contig of 3431 bp in length  
 30795: gap of unknown length  
 33519: contig of 2724 bp in length  
 33619: gap of unknown length  
 38662: contig of 5043 bp in length  
 38762: gap of unknown length  
 42466: contig of 3704 bp in length  
 42566: gap of unknown length  
 46288: contig of 4262 bp in length  
 46928: gap of unknown length  
 50787: contig of 3859 bp in length  
 50887: gap of unknown length  
 54203: contig of 3316 bp in length  
 54303: gap of unknown length  
 57668: contig of 3365 bp in length  
 57768: gap of unknown length  
 60823: contig of 3055 bp in length  
 60923: gap of unknown length  
 65117: contig of 4194 bp in length  
 65217: gap of unknown length  
 70453: contig of 5236 bp in length  
 70553: gap of unknown length  
 75210: contig of 4657 bp in length  
 75310: gap of unknown length  
 79469: contig of 4159 bp in length  
 79569: gap of unknown length  
 85488: contig of 5919 bp in length  
 85588: gap of unknown length  
 92583: contig of 6995 bp in length  
 92683: gap of unknown length  
 98505: contig of 5822 bp in length  
 98605: gap of unknown length  
 104564: contig of 5959 bp in length  
 104664: gap of unknown length  
 115203: contig of 10539 bp in length

```

* 115204 115303: gap of unknown length
* 115304 128133: contig of 12830 bp in length.
FEATURES
  Location/Qualifiers
    1..128133
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="2"
      /clone="RP11-438012"
    1..770
      /note="assembly_name:Contig18"
    871..2115
      /note="assembly_name:Contig28"
    2216..3046
      /note="assembly_name:Contig36"
    3147..4024
      /note="assembly_name:Contig37"
    4125..5010
      /note="assembly_name:Contig38"
    5111..5447
      /note="assembly_name:Contig39"
    5548..7913
      /note="assembly_name:Contig40"
    8014..9659
      /note="assembly_name:Contig42"
    9760..11026
      /note="assembly_name:Contig43"
    11127..12216
      /note="assembly_name:Contig44"
    12317..12509
      /note="assembly_name:Contig45"
    12610..14475
      /note="assembly_name:Contig46"
    14576..16865
      /note="assembly_name:Contig47"
    16966..19653
      /note="assembly_name:Contig48"
    19754..21718
      /note="assembly_name:Contig49"
    21819..24048
      /note="assembly_name:Contig50"
    24149..27164
      /note="assembly_name:Contig51"
    27265..30695
      /note="assembly_name:Contig52"
    30796..33519
      /note="assembly_name:Contig53"
    33620..38662
      /note="assembly_name:Contig54"
    38763..42466
      /note="assembly_name:Contig55"
    42567..46828
      /note="assembly_name:Contig56"
    46929..50787
      /note="assembly_name:Contig57"
    50888..54203
      /note="assembly_name:Contig58"
    54304..57668
      /note="assembly_name:Contig59"
    57769..60823
      /note="assembly_name:Contig60"
    60924..65117
      /note="assembly_name:Contig61"
    65218..70453
      /note="assembly_name:Contig62"
    clone_end:SP6
    vector_side:right
    70554..75210
      /note="assembly_name:Contig63"
    75311..79469
      /note="assembly_name:Contig64"
    79570..85488
      /note="assembly_name:Contig65"
    85589..92583

```

```

misc_feature      /note="assembly_name:Contig66"
92884..98505
misc_feature      /note="assembly_name:Contig67"
98606..104564
misc_feature      /note="assembly_name:Contig68"
104665..115203
misc_feature      /note="assembly_name:Contig69"
clone_end:T7

Query Match      1.7%  Score 19  DB 63  Length 128133;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 897 CTGCTGCTGAACCCCTGTG 915
|||||
Db 65300 CTGCTGCTGAACCCCTGTG 65318

RESULT 46
AC021488
LOCUS
DEFINITION
  AC021488 134738 bp DNA HTG 16-MAR-2000
  Homo sapiens chromosome 11 clone RP11-77F24 map 11, WORKING DRAFT
  SEQUENCE, 21 unordered pieces.
ACCESSION
  AC021488
VERSION
  AC021488.3 GI:7249184
KEYWORDS
  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 134738)
  Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  Homo sapiens chromosome 11, clone RP11-77F24
  Unpublished
REFERENCE
  2 (bases 1 to 134738)
  Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
  Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
  Boguslavsky,L., Boukhgaiter,B., Brown,A., Burkett,G., Castle,A.,
  Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
  DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
  Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
  Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
  Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
  Landers,T., Lechoczky,J., Levine,R., Liu,C., Liu,G., Locke,K.,
  Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
  McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
  Norman,C.H., O'Connor,T., O'Donnell,P., Olivat,T.M., Peterson,K.,
  Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
  Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
  Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
  Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
  Zimmer,A. and Zody,M.
  Direct Submission
TITLE
  JOURNAL
COMMENT
  Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  On Mar 16, 2000 this sequence version replaced gi:6850397.
  All repeats were identified using RepeatMasker:
  Smit, A.F.A. & Green, P. (1996-1997)
  http://ftp.genome.washington.edu/RM/RepeatMasker.html
  ----- Genome Center
  Center: Whitehead Institute/ MIT Center for Genome Research
  Center code: WIBR
  Web site: http://www-seq.wi.mit.edu
  Contact: sequence_submissions@genome.wi.mit.edu
  ----- Project Information
  Center project name: L5356
  Center clone name: 77_F24
  ----- Summary Statistics
  Sequencing vector: M13; M77815; 100% of reads
  Chemistry: Dye-terminator Big Dye; 100% of reads
  Assembly program: Phrap; version 0.960731
  Consensus quality: 124243 bases at least Q40
  Consensus quality: 129526 bases at least Q30

```

Consensus quality: 131528 bases at least Q20

Insert size: 139000; agarose-fp

Insert size: 132738; sum-of-contigs

Quality coverage: 3.5 in Q20 bases; agarose-fp

Quality coverage: 3.6 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 1559: contig of 1559 bp in length
* 1560 1659: gap of 100 bp
* 1660 2747: contig of 1088 bp in length
* 2748 2847: gap of 100 bp
* 2848 4474: contig of 1627 bp in length
* 4475 4574: gap of 100 bp
* 4575 6422: contig of 1848 bp in length
* 6423 6522: gap of 100 bp
* 6523 7638: contig of 1116 bp in length
* 7639 7738: gap of 100 bp
* 7739 9553: contig of 1815 bp in length
* 9554 9653: gap of 100 bp
* 9654 11602: contig of 1949 bp in length
* 11603 11702: gap of 100 bp
* 11703 15197: contig of 3495 bp in length
* 15198 15297: gap of 100 bp
* 15298 19894: contig of 4597 bp in length
* 19895 19994: gap of 100 bp
* 19995 23699: contig of 3705 bp in length
* 23700 23799: gap of 100 bp
* 23800 29731: contig of 5932 bp in length
* 29732 29831: gap of 100 bp
* 29832 32755: contig of 2924 bp in length
* 32756 32855: gap of 100 bp
* 32856 38266: contig of 5411 bp in length
* 38267 38366: gap of 100 bp
* 38367 45210: contig of 6844 bp in length
* 45211 45310: gap of 100 bp
* 45311 52507: contig of 7197 bp in length
* 52508 52607: gap of 100 bp
* 52608 59750: contig of 7143 bp in length
* 59751 59850: gap of 100 bp
* 59851 70139: contig of 10289 bp in length
* 70140 70239: gap of 100 bp
* 70240 83109: contig of 12870 bp in length
* 83110 83209: gap of 100 bp
* 83210 98621: contig of 15412 bp in length
* 98622 98721: gap of 100 bp
* 98722 112943: contig of 14222 bp in length
* 112944 113043: gap of 100 bp
* 113044 134738: contig of 21695 bp in length.
Location/Qualifiers
1. 134738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-77F24"
/clone_lib="RPC1-11 Human Male BAC"
1. 1559
/feature="assembly_fragment"
1660. 2747
/feature="assembly_fragment"
2848. 4474
/feature="assembly_fragment"
4575. 6422
/feature="assembly_fragment"
52508. 52607
/feature="assembly_fragment"
59751. 59850
/feature="assembly_fragment"
70140. 70239
/feature="assembly_fragment"
83110. 83209
/feature="assembly_fragment"
98622. 98721
/feature="assembly_fragment"
112944. 113043
/feature="assembly_fragment"
134738. 134738
/feature="assembly_fragment"

```

## FEATURES

source

```

misc_feature misc_end:T7
vector_side:left
7739. 9553
/feature="assembly_fragment"
9654. 11602
/feature="assembly_fragment"
11703. 15197
/feature="assembly_fragment"
15298. 19894
/feature="assembly_fragment"
19995. 23699
/feature="assembly_fragment"
23800. 29731
/feature="assembly_fragment"
29832. 32755
/feature="assembly_fragment"
32856. 38266
/feature="assembly_fragment"
38367. 45210
/feature="assembly_fragment"
45311. 52507
/feature="assembly_fragment"
52608. 59750
/feature="assembly_fragment"
59851. 70139
/feature="assembly_fragment"
70240. 83109
/feature="assembly_fragment"
clone_end:SP6
vector_side:left
83210. 98621
/feature="assembly_fragment"
98722. 112943
/feature="assembly_fragment"
113044. 134738
/feature="assembly_fragment"
BASE COUNT 42310 a 23650 c 25136 g 41636 t 2006 others
ORIGIN

Query Match 1.7%; Score 19; DB 67; Length 134738;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 940 TTCAGAAACTACTTTCTGA 958
|||||
Db 57205 TTCAGAAACTACTTTCTGA 57223

RESULT 47
AC011979 AC011979 137641 bp DNA HTG 30-MAR-2000
LOCUS Homo sapiens clone RP11-16F15, WORKING DRAFT SEQUENCE, 30 unordered
DEFINITION pieces.
AC011979
VERSION AC011979.3 GI:7341927
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 137641)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Allen,N., Anderson,M.,
TITLE Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
JOURNAL Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A.,
REFERENCE Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donellan,L., Doyle,M.,
AUTHORS Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

```

Lehoczyk, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,  
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,  
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (17-OCT-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 30, 2000 this sequence version replaced gi:6453971.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L3529

Center clone name: 16\_F\_15

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 121496 bases at least Q40

Consensus quality: 129060 bases at least Q30

Consensus quality: 132115 bases at least Q20

Insert size: 154000; agarose-fp

Insert size: 134741; sum-of-contigs

Quality coverage: 3.4 in Q20 bases; agarose-fp

Quality coverage: 3.9 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 30 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1 1611: contig of 1611 bp in length
* 1612 1711: gap of 100 bp
* 1712 3086: contig of 1375 bp in length
* 3087 3186: gap of 100 bp
* 3187 4590: contig of 1404 bp in length
* 4591 4690: gap of 100 bp
* 4691 5932: contig of 1242 bp in length
* 5933 6032: gap of 100 bp
* 6033 7558: contig of 1526 bp in length
* 7559 7658: gap of 100 bp
* 7659 9074: contig of 1416 bp in length
* 9075 9174: gap of 100 bp
* 9175 10813: contig of 1639 bp in length
* 10814 10913: gap of 100 bp
* 10914 12650: contig of 1737 bp in length
* 12651 12750: gap of 100 bp
* 12751 14999: contig of 2249 bp in length
* 15000 15099: gap of 100 bp
* 15100 17144: contig of 2045 bp in length
* 17145 17244: gap of 100 bp
* 17245 19475: contig of 2231 bp in length
* 19476 19575: gap of 100 bp
* 19576 22387: contig of 2812 bp in length
* 22388 22487: gap of 100 bp
* 22488 22525: contig of 788 bp in length
* 22526 23353: gap of 100 bp
* 23356 26261: contig of 2906 bp in length
* 26262 26361: gap of 100 bp
* 26362 29293: contig of 2932 bp in length
* 29294 29393: gap of 100 bp
* 29394 33323: contig of 3930 bp in length

```

```

* 33324 33423: gap of 100 bp
* 33424 38990: contig of 5567 bp in length
* 38991 39090: gap of 100 bp
* 39091 43165: contig of 4075 bp in length
* 43166 43265: gap of 100 bp
* 43266 48702: contig of 5437 bp in length
* 48703 48802: gap of 100 bp
* 48803 53681: contig of 4879 bp in length
* 53682 53781: gap of 100 bp
* 53782 58325: contig of 4544 bp in length
* 58326 58425: gap of 100 bp
* 58426 64502: contig of 6077 bp in length
* 64503 64602: gap of 100 bp
* 64603 72223: contig of 7621 bp in length
* 72224 72323: gap of 100 bp
* 72324 79639: contig of 7316 bp in length
* 79640 79739: gap of 100 bp
* 79740 88137: contig of 8398 bp in length
* 88138 88237: gap of 100 bp
* 88238 96553: contig of 8316 bp in length
* 96554 96653: gap of 100 bp
* 96654 105243: contig of 8590 bp in length
* 105244 105343: gap of 100 bp
* 105344 113380: contig of 8037 bp in length
* 113381 113480: gap of 100 bp
* 113481 122965: contig of 9485 bp in length
* 122966 123065: gap of 100 bp
* 123066 137641: contig of 14576 bp in length.

```

#### FEATURES

##### source

```

1..137641
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RP11-16F15"
/clone_lib="RPC1-11 Human Male BAC"

```

```

misc_feature      1..1611
                  /note="assembly_fragment"
misc_feature      1712..3086
                  /note="assembly_fragment"
misc_feature      3187..4590
                  /note="assembly_fragment"
misc_feature      4691..5932
                  /note="assembly_fragment"
misc_feature      6033..7558
                  /note="assembly_fragment"
misc_feature      7659..9074
                  /note="assembly_fragment"
misc_feature      9175..10813
                  /note="assembly_fragment"
misc_feature      10914..12650
                  /note="assembly_fragment"
misc_feature      12751..14999
                  /note="assembly_fragment"
misc_feature      15100..17144
                  /note="assembly_fragment"
misc_feature      17245..19475
                  /note="assembly_fragment"
misc_feature      19576..22387
                  /note="assembly_fragment"
misc_feature      22488..23255
                  /note="assembly_fragment"
                  clone_end:SP6
                  vector_side:right
misc_feature      23356..26261
                  /note="assembly_fragment"
misc_feature      26362..29293
                  /note="assembly_fragment"
misc_feature      29394..33323
                  /note="assembly_fragment"
misc_feature      33424..38990
                  /note="assembly_fragment"
misc_feature      39091..43165
                  /note="assembly_fragment"
misc_feature      43266..48702

```

```

/note="assembly_fragment"
48803..53681
/note="assembly_fragment"
53782..58325
/note="assembly_fragment"
58426..64502
/note="assembly_fragment"
64603..72223
/note="assembly_fragment"
72324..79639
/note="assembly_fragment"
79740..88137
/note="assembly_fragment"
88238..96553
/note="assembly_fragment"
96654..105243
/note="assembly_fragment"
105344..113380
/note="assembly_fragment"
113481..122965
/note="assembly_fragment"
clone_end:17
vector_side:right"
123066..137641
/note="assembly_fragment"
39168 a 29416 c 27916 g 38239 t 2902 others
ORIGIN

Query Match      1.7%  Score 19;  DB 62;  Length 137641;
Best Local Similarity 100.0%;  Pred. No. 45;
Matches 19;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 158 TGGGAATATCTGGTGGT 176
|||||
Db 61244 TGGGAATATCTGGTGGT 61262

RESULT 48
AC005845
LOCUS
DEFINITION
Homo sapiens chromosome 12 clone RP3-454B23, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
AC005845
VERSION
AC005845.5 GI:11094658
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 142530)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Kratison,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulisege,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

```

Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
Ogoh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojebokan,I., Rolfe,M.,  
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,  
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,  
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,  
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
and Gibbs,R.

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: ZR  
Center clone name: RP3-454B23  
----- Summary Statistics  
Sequencing vector: Plasmid; M77789  
Chemistry: Dye-terminator Big Dye; 26% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 139433 bases at least Q40  
Consensus quality: 143241 bases at least Q30  
Consensus quality: 145391 bases at least Q20  
Estimated insert size: 137681; sum-of-contigs estimation  
Quality coverage: 11.7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 58734: contig of 58734 bp in length  
\* 58735 58834: gap of unknown length  
\* 58835 97561: contig of 38727 bp in length  
\* 97562 97661: gap of unknown length  
\* 97662 125825: contig of 28164 bp in length  
\* 125826 125925: gap of unknown length  
\* 125926 136855: contig of 10930 bp in length  
\* 136856 136955: gap of unknown length  
\* 136956 138475: contig of 1520 bp in length  
\* 138476 138575: gap of unknown length  
\* 138576 140084: contig of 1509 bp in length  
\* 140085 140184: gap of unknown length  
\* 140185 141333: contig of 1149 bp in length  
\* 141334 141433: gap of unknown length  
\* 141434 142530: contig of 1097 bp in length.

Location/Qualifiers  
1..142530  
/organism="Homo sapiens"

FEATURES  
source

```

/db_xref="taxon:9606"
/clone="RP3-454B23"
BASE COUNT 38009 a 34230 c 34450 g 34705 t 1136 others
ORIGIN

Query Match 1.7%; Score 19; DB 60; Length 142530;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 TGCCAGTGCAGTGTGCTA 336
|||||
Db 128842 TGCCAGTGCAGTGTGCTA 128860

RESULT 49
AC009607/c
LOCUS
DEFINITION Homo sapiens clone RP11-1111, WORKING DRAFT SEQUENCE, 31 unordered
pieces.
ACCESSION AC009607
VERSION AC009607.3 GI:8072446
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 146055)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-1111
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 146055)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Testaye,S., Toriella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7321520.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L1461
Center clone name: L111
----- Summary Statistics
Sequencing vector: M13; M7815; 99% of reads
Sequencing vector: Plasmid; n/a; 0.0% of reads
Chemistry: Dye-terminator; Dye-primer-amersham; 6% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 126247 bases at least Q40
Consensus quality: 134710 bases at least Q30
Consensus quality: 138332 bases at least Q20
Insert size: 160000; agarose-fp
FEATURES
source
1. .146055
Insert size: 143055; sum-of-contigs
Quality coverage.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1166: contig of 1166 bp in length
1167 1266: gap of 100 bp
1267 2417: contig of 1151 bp in length
2418 2517: gap of 100 bp
2518 3635: contig of 1118 bp in length
3636 3735: gap of 100 bp
3736 4851: contig of 1116 bp in length
4852 4951: gap of 100 bp
4952 6701: contig of 1750 bp in length
6702 6801: gap of 100 bp
6802 8629: contig of 1828 bp in length
8630 8729: gap of 100 bp
8730 10482: contig of 1753 bp in length
10483 10582: gap of 100 bp
10583 12126: contig of 1544 bp in length
12127 12226: gap of 100 bp
12227 13831: contig of 1605 bp in length
13832 13931: gap of 100 bp
13932 16132: contig of 2201 bp in length
16133 16232: gap of 100 bp
16233 17951: contig of 1719 bp in length
17952 18051: gap of 100 bp
18052 20069: contig of 2018 bp in length
20070 20169: gap of 100 bp
20170 22565: contig of 2396 bp in length
22566 23665: gap of 100 bp
23666 24924: contig of 2259 bp in length
24925 25024: gap of 100 bp
25025 26996: contig of 1972 bp in length
26997 27096: gap of 100 bp
27097 28834: contig of 1738 bp in length
28835 28934: gap of 100 bp
28935 33803: contig of 4869 bp in length
33804 33903: gap of 100 bp
33904 38177: contig of 4274 bp in length
38178 38277: gap of 100 bp
38278 42132: contig of 3855 bp in length
42133 42232: gap of 100 bp
42233 45874: contig of 3642 bp in length
45875 45974: gap of 100 bp
45975 50807: contig of 4833 bp in length
50808 50907: gap of 100 bp
50909 55690: contig of 4783 bp in length
55691 55790: gap of 100 bp
55791 60868: contig of 5078 bp in length
60869 60968: gap of 100 bp
60969 65180: contig of 4212 bp in length
65181 65280: gap of 100 bp
65281 72495: contig of 7215 bp in length
72496 72595: gap of 100 bp
72596 79536: contig of 6941 bp in length
79537 79636: gap of 100 bp
79637 86013: contig of 6377 bp in length
86014 86113: gap of 100 bp
86114 98467: contig of 12354 bp in length
98468 98567: gap of 100 bp
98568 112312: contig of 13745 bp in length
112313 112412: gap of 100 bp
112413 128733: contig of 16321 bp in length
128734 128833: gap of 100 bp
128834 146055: contig of 17222 bp in length.
Location/Qualifiers
1. .146055

```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RP11-1111"
/clone_lib="RP11-1111" Human Male BAC"
1..1166
/note="assembly_fragment"
1267..2417
/note="assembly_fragment"
2518..3635
/note="assembly_fragment"
3736..4851
/note="assembly_fragment"
4952..6701
/note="assembly_fragment"
6802..8629
/note="assembly_fragment"
8730..10482
/note="assembly_fragment"
10583..12126
/note="assembly_fragment"
12227..13831
/note="assembly_fragment"
13932..16132
/note="assembly_fragment"
16233..17951
/note="assembly_fragment"
18052..20069
/note="assembly_fragment"
20170..22565
/note="assembly_fragment"
22666..24924
/note="assembly_fragment"
25025..26996
/note="assembly_fragment"
27097..28834
/note="assembly_fragment
clone_end:SP6
vector_side:left"
28935..33803
/note="assembly_fragment"
33904..38177
/note="assembly_fragment"
38278..42132
/note="assembly_fragment"
42233..45874
/note="assembly_fragment
clone_end:T7
vector_side:left"
45975..50807
/note="assembly_fragment"
50908..55690
/note="assembly_fragment"
55791..60868
/note="assembly_fragment"
60969..65180
/note="assembly_fragment"
65281..72495
/note="assembly_fragment"
72596..79536
/note="assembly_fragment"
79637..86013
/note="assembly_fragment"
86114..98467
/note="assembly_fragment"
98568..112312
/note="assembly_fragment"
112413..128733
/note="assembly_fragment"
128834..146055
/note="assembly_fragment"
41047 a 30113 c 30523 g 41366 t 3006 others
ORIGIN
```

```
Query Match 1.7%; Score 19; DB 61; Length 146055;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 TCACAGAGTCCTGGCTTT 890
|||||
Db 12704 TCACAGAGTCCTGGCTTT 12686

RESULT 50
AF188027/c
LOCUS
DEFINITION
Homo sapiens chromosome 8 map 8q21 clone RP11-101A8, WORKING DRAFT
SEQUENCE, 2 ordered pieces.
ACCESSION AF188027
VERSION AF188027.2 GI:8152141
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 146511)
Taudien,S., Blechschmidt,K., Menzel,U., Polley,A., Reichwald,K.,
Rump,A., Schilhabel,M.B., Schudy,A., Wen,G. and Rosenthal,A.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 146511)
Taudien,S., Wen,G.P., Schilhabel,M., Menzel,U., Jahn,N.,
Baumgart,C., Dettie,M. and Rosenthal,A.
Direct Submission
Submitted (09-SEP-1999) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
3 (bases 1 to 146511)
Genome Sequencing Center Jena.
Direct Submission
Submitted (24-MAY-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Jun 1, 2000 this sequence version replaced gi:6007576.
* NOTE: This is a 'working draft' sequence..it currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 106240: contig of 106240 bp in length
* 106241 106340: gap of unknown length
* 106341 146511: contig of 40171 bp in length.
```

## FEATURES

## source

```
1..146511
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8 map 8q21"
/clone="RP11-101A8"
```

## misc\_feature

```
1..106240
/note="assembly_fragment
clone_end:SP6
vector_side:left"
```

## misc\_feature

```
106341..146511
/note="assembly_fragment
clone_end:T7
vector_side:right"
```

```
BASE COUNT 43231 a 29758 c 30108 g 43313 t 101 others
ORIGIN
```

```
Query Match 1.7%; Score 19; DB 78; Length 146511;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 811 AATTTGGTAAATGAACC 829
```

|||||  
Db 86767 AATTGGGTAAATGAACC 86749

Search completed: May 24, 2001, 09:32:41  
Job time: 5617 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2001, 07:46:29 ; Search time 1187.17 Seconds  
(without alignments)  
8234.452 Million cell updates/sec

Title: US-08-887-977-9  
Perfect score: 1119  
Sequence: 1 ATGTTTTCAGTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 9623517 seqs, 4368049070 residues

Word size : 15  
Total number of hits satisfying chosen parameters: 12826

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:\*

- 1: gb\_estl1.\*
- 2: gb\_estl2.\*
- 3: gb\_estl3.\*
- 4: gb\_estl4.\*
- 5: gb\_estl5.\*
- 6: gb\_estl6.\*
- 7: gb\_estl7.\*
- 8: gb\_estl8.\*
- 9: gb\_estl9.\*
- 10: gb\_estl10.\*
- 11: gb\_estl11.\*
- 12: gb\_estl12.\*
- 13: gb\_estl13.\*
- 14: gb\_estl14.\*
- 15: gb\_estl15.\*
- 16: gb\_estl16.\*
- 17: gb\_estl17.\*
- 18: gb\_estl18.\*
- 19: gb\_estl19.\*
- 20: gb\_estl20.\*
- 21: gb\_estl21.\*
- 22: gb\_estl22.\*
- 23: gb\_estl23.\*
- 24: gb\_estl24.\*
- 25: gb\_estl33.\*
- 26: gb\_estl34.\*
- 27: gb\_estl35.\*
- 28: gb\_estl36.\*
- 29: gb\_estl37.\*
- 30: gb\_estl38.\*
- 31: gb\_estl39.\*
- 32: gb\_estl40.\*
- 33: em\_estba.\*
- 34: em\_estfun.\*
- 35: em\_esthum1.\*
- 36: em\_esthum2.\*
- 37: em\_esthum3.\*
- 38: em\_esthum4.\*
- 39: em\_esthum5.\*
- 40: em\_esthum6.\*
- 41: em\_esthum7.\*
- 42: em\_esthum8.\*
- 43: em\_esthum9.\*

- 44: em\_esthum10.\*
- 45: em\_esthum11.\*
- 46: em\_esthum12.\*
- 47: em\_esthum13.\*
- 48: em\_esthum14.\*
- 49: em\_esthum15.\*
- 50: em\_esthum16.\*
- 51: em\_esthum17.\*
- 52: em\_esthum18.\*
- 53: em\_esthum19.\*
- 54: em\_esthum20.\*
- 55: em\_esthum21.\*
- 56: em\_esthum22.\*
- 57: em\_esthum23.\*
- 58: em\_esthum24.\*
- 59: em\_esthum25.\*
- 60: em\_esthum26.\*
- 61: em\_esthum27.\*
- 62: em\_esthum28.\*
- 63: em\_estin1.\*
- 64: em\_estin2.\*
- 65: em\_estin3.\*
- 66: em\_estin4.\*
- 67: em\_estin5.\*
- 68: em\_estomi.\*
- 69: em\_estom2.\*
- 70: em\_estov1.\*
- 71: em\_estov2.\*
- 72: em\_estpl1.\*
- 73: em\_estpl2.\*
- 74: em\_estpl3.\*
- 75: em\_estpl4.\*
- 76: em\_estpl5.\*
- 77: em\_estpl6.\*
- 78: em\_estpl7.\*
- 79: em\_estpl8.\*
- 80: em\_estpl9.\*
- 81: em\_estpl10.\*
- 82: em\_estro1.\*
- 83: em\_estro2.\*
- 84: em\_estro3.\*
- 85: em\_estro4.\*
- 86: em\_estro5.\*
- 87: em\_estro6.\*
- 88: em\_estro7.\*
- 89: em\_estro8.\*
- 90: em\_estro9.\*
- 91: em\_estro10.\*
- 92: em\_estro11.\*
- 93: em\_estro12.\*
- 94: em\_estro13.\*
- 95: em\_estro14.\*
- 96: em\_estro15.\*
- 97: em\_estro16.\*
- 98: em\_estro17.\*
- 99: em\_estro18.\*
- 100: em\_estro19.\*
- 101: em\_estro20.\*
- 102: gb\_est25.\*
- 103: gb\_est26.\*
- 104: gb\_est27.\*
- 105: gb\_est28.\*
- 106: gb\_est29.\*
- 107: gb\_est30.\*
- 108: gb\_est31.\*
- 109: gb\_est32.\*
- 110: gb\_est41.\*
- 111: gb\_est42.\*
- 112: gb\_est43.\*
- 113: gb\_est44.\*
- 114: gb\_est45.\*
- 115: gb\_est46.\*
- 116: gb\_est47.\*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
```

```
190: em_gss_pln1:*
191: em_gss_pln2:*
192: em_gss_pro:*
193: em_gss_rod1:*
194: em_gss_rod2:*
195: em_gss_rod3:*
196: em_gss_rod4:*
197: em_gss_rod5:*
198: em_gss_vrt1:*
199: em_gss_vrt2:*
200: em_gss_vrt3:*
201: gb_gss1:*
202: gb_gss2:*
203: gb_gss3:*
204: gb_gss4:*
205: gb_gss5:*
206: gb_gss6:*
207: gb_gss7:*
208: gb_gss8:*
209: gb_gss9:*
210: gb_gss10:*
211: gb_gss11:*
212: gb_gss12:*
213: gb_gss13:*
214: gb_gss14:*
215: gb_gss15:*
216: gb_gss16:*
217: gb_gss17:*
218: gb_gss18:*
219: gb_gss19:*
220: gb_gss20:*
221: gb_gss21:*
222: gb_gss22:*
223: gb_gss23:*
224: gb_gss24:*
225: gb_gss25:*
226: gb_gss26:*
227: gb_gss27:*
228: gb_gss28:*
229: gb_gss29:*
230: gb_gss30:*
231: gb_gss31:*
232: gb_gss32:*
233: gb_gss33:*
234: gb_gss34:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description	
1	114	10.2	522	120	AW855262	AW855262	PM1-CT026	
C	2	79	7.1	382	AI808029	AI808029	wf52904.x	
C	3	23	2.1	461	AI045155	AI045155	UI-R-C1-k	
4	21	1.9	822	144	BF105117	BF105117	601822118	
5	20	1.8	396	141	BE950816	BE950816	UI-M-CEO-	
C	6	20	1.8	412	N83058	N83058	TgESTzy59g1	
7	20	1.8	489	172	BG044568	BG044568	saa30a06.	
C	8	20	1.8	513	AZ023717	AZ023717	RPCI-23-3	
C	9	20	1.8	530	AW733824	AW733824	sk84h02.y	
10	20	1.8	568	119	AW773091	AW773091	925002C05	
C	11	20	1.8	609	AZ695745	AZ695745	RPCI-23-2	
12	20	1.8	731	230	CNS02TWL	AL213726	Tetraodon	
13	19	1.7	260	138	BE698296	BE698296	RC2-UT002	
C	14	19	1.7	262	206	AQ413053	AQ413053	RPCI-11-1
C	15	19	1.7	276	4	AA245646	AA245646	mx01a03.r
16	19	1.7	280	209	AQ639171	AQ639171	927P1-1H5	
17	19	1.7	330	19	AI374108	AI374108	T6314 MVA	
18	19	1.7	348	19	AI374416	AI374416	T6773 MVA	

19	1.7	355	24	A1707310	T7429 MVA	18	1.6	463	4	AA225058	nc21b12.r
20	1.7	360	20	A1464788	mz80h09.y	18	1.6	463	207	AQ466845	HS_5188_A
21	1.7	360	158	W84101	T2925 MVA74	18	1.6	473	168	BF773895	283409 MA
22	1.7	362	19	A1374268	T6480 MVA	18	1.6	478	203	AQ224100	HS_2011_A
23	1.7	367	17	A1215388	T6273 MVA	18	1.6	500	21	A1533590	SD05477.5
24	1.7	370	24	A1707317	T7437 MVA	18	1.6	507	142	BF006661	ESP435159
25	1.7	402	21	A1510879	T6885 MVA	18	1.6	509	8	AA531233	nj52606.s
26	1.7	435	1	AA014588	mg93a01.r	18	1.6	509	203	AQ227870	HS_2020_B
27	1.7	455	233	TA390D01Q		18	1.6	514	130	AQ720078	HS_5542_A
28	1.7	457	4	AA254935	AL498801 T. brucei	18	1.6	518	137	BE638669	946012E02
29	1.7	463	4	AA254935	AA254935 mz80h09.r	18	1.6	520	206	AQ417897	RPCI-11-1
30	1.7	468	29	AV397592	AV397592 mz82f03.f	18	1.6	530	225	AZ654674	IM0528P21
31	1.7	481	103	A1881128	A1881128 T8077 MVA	18	1.6	538	118	AW629816	hh71c05.y
32	1.7	484	148	BF443898	BF443898 T61560 MA	18	1.6	544	202	AQ118433	HS_3007_A
33	1.7	490	225	AZ646557	AZ646557 1M0512F11	18	1.6	545	209	AQ156432	Sheared_D
34	1.7	494	21	A1552452	A1552452 mg93a01.y	18	1.6	545	210	AQ734398	HS_3012_A
35	1.7	502	1	AA060911	AA060911 mj88b01.f	18	1.6	557	207	AQ79020	RPCI-11-2
36	1.7	506	201	AQ003850	AQ003850 CIT-HSP-2	18	1.6	558	172	BG052009	RH122-7_G
37	1.7	519	103	A1880998	A1880998 T7766 MVA	18	1.6	565	210	AQ726314	HS_5406_B
38	1.7	534	215	AZ016585	AZ016585 RPCI-23-3	18	1.6	569	220	AZ382920	1M0140F22
39	1.7	562	161	BE033064	BE033064 133359 MA	18	1.6	570	205	AQ384598	RPCI11-14
40	1.7	568	161	BE033241	BE033241 133608 MA	18	1.6	571	204	AQ309108	CIT-HSP-2
41	1.7	576	3	AA148128	AA148128 z031a06.r	18	1.6	580	218	AZ248187	RPCI-23-1
42	1.7	588	167	BE468435	BE468435 IPHdk0265	18	1.6	584	121	AW864107	PMO-SN001
43	1.7	626	122	AW956448	AW956448 EST368518	18	1.6	584	164	BE216221	HV-CEB000
44	1.7	631	151	BF660122	BF660122 ma27h06.	18	1.6	586	164	BE249133	NFO38C11D
45	1.7	633	214	AZ003038	AZ003038 RPCI-23-3	18	1.6	586	209	AQ672695	HS_2122_A
46	1.7	640	205	AQ389111	AQ389111 RPCI11-15	18	1.6	588	207	AQ523699	HS_5221_A
47	1.7	724	205	AQ379360	AQ379360 RPCI11-16	18	1.6	596	223	AZ514095	IM0360J08
48	1.7	785	15	A1047830	A1047830 ud64b12.x	18	1.6	600	208	AQ606976	HS_5397_B
49	1.7	788	145	BF235138	BF235138 602026409	18	1.6	632	33	A1647936	uk38a02.x
50	1.7	802	20	A1430701	A1430701 mc59h06.y	18	1.6	639	31	AV653345	AV653345
51	1.7	834	139	BE777635	BE777635 6013348368	18	1.6	653	32	AV752881	AV752881
52	1.7	868	230	CNS022RE	AL221315 Tetraodon	18	1.6	654	145	BF245892	NO1864070
53	1.7	1030	231	CNS03Q9B	AL256565 Tetraodon	18	1.6	666	151	BF632993	FP048E02D
54	1.7	1036	230	CNS02KMI	AL201682 Tetraodon	18	1.6	666	166	BE267868	601125412
55	1.7	1098	230	CNS02KY9	AL202122 Tetraodon	18	1.6	681	207	AQ477473	CITBI-E1-
56	1.6	127	13	AA919945	AA919945 vy47g04.r	18	1.6	689	143	BF034701	601455565
57	1.6	173	211	AQ979419	AQ979419 Y35U CEPH	18	1.6	749	149	BF525475	602069564
58	1.6	190	223	AZ517092	AZ517092 RPCI-11-8	18	1.6	771	150	BF575792	602134829
59	1.6	228	103	A1904352	A1904352 IL-BT051-	18	1.6	782	29	AV404402	AV404402
60	1.6	248	120	AW846966	AW846966 RCI-CT019	18	1.6	823	166	BE410806	601301442
61	1.6	260	29	AV379721	AV379721 AV379721	18	1.6	831	217	AZ184799	SP-1003_B
62	1.6	273	174	BG235359	BG235359 NCEST3a39	18	1.6	831	231	CNS03QPN	Tetraodon
63	1.6	276	133	BB437562	BB437562 BB437562	18	1.6	878	204	AQ259828	nbxb00240
64	1.6	282	126	BB140981	BB140981 BB140981	18	1.6	956	173	BG113517	602283975
65	1.6	285	109	AV116567	AV116567 AV116567	18	1.6	966	231	CNS043MN	Tetraodon
66	1.6	289	136	BE522281	BE522281 M24G8XTM	18	1.6	970	217	AZ166622	SP-0088_B
67	1.6	307	125	BB105591	BB105591 BB105591	18	1.6	989	139	BE798348	601582103
68	1.6	311	159	Z35153	Z35153 ATTS3711 St	18	1.6	1010	230	CNS0216F	Tetraodon
69	1.6	344	174	BG147142	BG147142 mab48R03	18	1.6	1222	149	BF527833	602040947
70	1.6	357	203	AQ193807	AQ193807 CIT-HSP-2	18	1.6	1343	174	BG169441	602321127
71	1.6	372	210	AQ695729	AQ695729 HS_5491_B	18	1.6	96	219	AZ310465	IM0025L23
72	1.6	381	227	B31064	B31064 HS-1005-B1-	17	1.5	136	218	AZ246625	RPCI-23-9
73	1.6	386	202	AQ133524	AQ133524 HS_3051_A	17	1.5	163	3	AA190203	TH058 HTC
74	1.6	386	203	AQ185889	AQ185889 HS_2240_A	17	1.5	175	152	D76737	MUS70B09 mo
75	1.6	399	162	BE102943	BE102943 UI-R-BT1-	17	1.5	181	112	AW236904	xm50f11.x
76	1.6	401	205	AQ344425	AQ344425 RPCI11-11	17	1.5	200	162	BE061217	CM1-BR005
77	1.6	403	156	R83706	R83706 yq14e10.r1	17	1.5	205	117	AW567540	x28h09.x
78	1.6	415	156	R95058	R95058 yq44f03.r1	17	1.5	208	147	BF361456	CMO-PT005
79	1.6	418	151	BF634199	BF634199 NF087H11D	17	1.5	211	20	AT423838	tg35b11.x
80	1.6	420	162	BE099959	BE099959 UI-R-BJ1-	17	1.5	215	147	BF364851	MR3-NN112
81	1.6	423	227	B53590	B53590 CIT-HSP-201	17	1.5	219	147	BF360309	MR0-MR012
82	1.6	427	207	AQ469740	AQ469740 CITBI-EI-	17	1.5	229	168	BF701962	MI-P-E3-a
83	1.6	429	212	AQ058338	AQ058338 nbe600120	17	1.5	233	123	BB033977	BB033977
84	1.6	431	151	BF632538	BF632538 ov59b10.s	17	1.5	243	32	AW003686	wq82d12.x
85	1.6	437	222	AZ482104	AZ482104 1M030710A	17	1.5	244	117	AW602914	CMO-BT060
86	1.6	450	112	AW243929	AW243929 x058g02.x	17	1.5	246	107	A1217207	qf73c08.x
87	1.6	450	115	AW450179	AW450179 UI-H-BI3-	17	1.5	248	103	A1863545	wh72g06.x
88	1.6	450	208	AQ056946	AQ056946 HS_5328_A	17	1.5	253	158	W13520	mag4c03.r1
89	1.6	461	155	R36281	R36281 y968c11.r1	17	1.5	254	143	BF056596	7k20a03.x
90	1.6	461	155	R36281	W25156 zb69c07.r1	17	1.5	262	8	AA495129	fa03c04.r
91	1.6	462	158	W25156		17	1.5				

165	17	1.5	264	8	AA492112	AA492112 ngl7d01.s	C 238	17	1.5	376	226	AZ699465
166	17	1.5	265	219	AZ312321	AZ312321 1M0028A08	C 239	17	1.5	376	227	B34293
167	17	1.5	266	4	AA243629	AA243629 zs16d09.s	C 240	17	1.5	377	201	AQ042822
168	17	1.5	267	147	BF366711	BF366711 IL3-NF010	C 241	17	1.5	378	9	AA568772
169	17	1.5	268	220	AZ366760	AZ366760 1M0116F08	C 242	17	1.5	379	141	BE947763
170	17	1.5	269	161	BE014055	BE014055 125631 MA	C 243	17	1.5	379	155	N81275
171	17	1.5	270	160	BE573231	BE573231 BB573231	C 244	17	1.5	381	22	AI587014
172	17	1.5	271	161	BB080676	BB080676 BB080676	C 245	17	1.5	381	214	AQ965702
173	17	1.5	272	128	BB250477	BB250477 BB250477	C 246	17	1.5	382	150	BF554147
174	17	1.5	273	128	BB111355	BB111355 BB111355	C 247	17	1.5	382	227	B79973
175	17	1.5	274	128	BB229347	BB229347 BB229347	C 248	17	1.5	383	213	AQ884619
176	17	1.5	275	131	BB355837	BB355837 BB355837	C 249	17	1.5	384	215	AI028168
177	17	1.5	276	131	BB355837	BB355837 BB355837	C 250	17	1.5	385	18	AI262771
178	17	1.5	277	134	BB466180	BB466180 BB466180	C 251	17	1.5	386	157	T64598
179	17	1.5	278	134	BB466180	BB466180 BB466180	C 252	17	1.5	388	136	BE547381
180	17	1.5	279	134	BB466180	BB466180 BB466180	C 253	17	1.5	389	3	AA199418
181	17	1.5	280	134	BB466180	BB466180 BB466180	C 254	17	1.5	390	143	BF092215
182	17	1.5	281	134	BB466180	BB466180 BB466180	C 255	17	1.5	391	32	AW016574
183	17	1.5	282	134	BB466180	BB466180 BB466180	C 256	17	1.5	391	173	BG132167
184	17	1.5	283	134	BB466180	BB466180 BB466180	C 257	17	1.5	392	121	AW860752
185	17	1.5	284	134	BB466180	BB466180 BB466180	C 258	17	1.5	393	19	AI401252
186	17	1.5	285	134	BB466180	BB466180 BB466180	C 259	17	1.5	393	116	AW524977
187	17	1.5	286	134	BB466180	BB466180 BB466180	C 260	17	1.5	395	214	AQ967189
188	17	1.5	287	134	BB466180	BB466180 BB466180	C 261	17	1.5	403	113	AQ310988
189	17	1.5	288	134	BB466180	BB466180 BB466180	C 262	17	1.5	403	116	AW486365
190	17	1.5	289	134	BB466180	BB466180 BB466180	C 263	17	1.5	404	103	AI872253
191	17	1.5	290	134	BB466180	BB466180 BB466180	C 264	17	1.5	405	159	Z46531
192	17	1.5	291	134	BB466180	BB466180 BB466180	C 265	17	1.5	406	165	BE333043
193	17	1.5	292	134	BB466180	BB466180 BB466180	C 266	17	1.5	407	141	BE921946
194	17	1.5	293	134	BB466180	BB466180 BB466180	C 267	17	1.5	408	3	AA218271
195	17	1.5	294	134	BB466180	BB466180 BB466180	C 268	17	1.5	409	144	BF151422
196	17	1.5	295	134	BB466180	BB466180 BB466180	C 269	17	1.5	409	162	BE097602
197	17	1.5	296	134	BB466180	BB466180 BB466180	C 270	17	1.5	413	13	AA914068
198	17	1.5	297	134	BB466180	BB466180 BB466180	C 271	17	1.5	413	9	AA593453
199	17	1.5	298	134	BB466180	BB466180 BB466180	C 272	17	1.5	413	115	AW433323
200	17	1.5	299	134	BB466180	BB466180 BB466180	C 273	17	1.5	414	227	B30657
201	17	1.5	300	134	BB466180	BB466180 BB466180	C 274	17	1.5	415	229	CNS00RZF
202	17	1.5	301	134	BB466180	BB466180 BB466180	C 275	17	1.5	417	13	AA914068
203	17	1.5	302	134	BB466180	BB466180 BB466180	C 276	17	1.5	417	13	AI060032
204	17	1.5	303	134	BB466180	BB466180 BB466180	C 277	17	1.5	417	30	AV561957
205	17	1.5	304	134	BB466180	BB466180 BB466180	C 278	17	1.5	419	158	W66400
206	17	1.5	305	134	BB466180	BB466180 BB466180	C 279	17	1.5	421	167	BE494624
207	17	1.5	306	134	BB466180	BB466180 BB466180	C 280	17	1.5	421	203	AQ206723
208	17	1.5	307	134	BB466180	BB466180 BB466180	C 281	17	1.5	421	213	AQ937198
209	17	1.5	308	134	BB466180	BB466180 BB466180	C 282	17	1.5	422	106	AI889610
210	17	1.5	309	134	BB466180	BB466180 BB466180	C 283	17	1.5	422	116	AW524292
211	17	1.5	310	134	BB466180	BB466180 BB466180	C 284	17	1.5	422	122	AW990203
212	17	1.5	311	134	BB466180	BB466180 BB466180	C 285	17	1.5	422	163	BE159985
213	17	1.5	312	134	BB466180	BB466180 BB466180	C 286	17	1.5	422	163	BE160028
214	17	1.5	313	134	BB466180	BB466180 BB466180	C 287	17	1.5	423	18	AI255756
215	17	1.5	314	134	BB466180	BB466180 BB466180	C 288	17	1.5	423	23	AI682005
216	17	1.5	315	134	BB466180	BB466180 BB466180	C 289	17	1.5	426	23	AI682005
217	17	1.5	316	134	BB466180	BB466180 BB466180	C 290	17	1.5	427	19	AI400540
218	17	1.5	317	134	BB466180	BB466180 BB466180	C 291	17	1.5	428	17	AI190134
219	17	1.5	318	134	BB466180	BB466180 BB466180	C 292	17	1.5	429	20	AI435212
220	17	1.5	319	134	BB466180	BB466180 BB466180	C 293	17	1.5	429	107	AU168422
221	17	1.5	320	134	BB466180	BB466180 BB466180	C 294	17	1.5	430	24	AI717948
222	17	1.5	321	134	BB466180	BB466180 BB466180	C 295	17	1.5	431	10	AA683382
223	17	1.5	322	134	BB466180	BB466180 BB466180	C 296	17	1.5	432	14	AA957786
224	17	1.5	323	134	BB466180	BB466180 BB466180	C 297	17	1.5	432	115	AW451401
225	17	1.5	324	134	BB466180	BB466180 BB466180	C 298	17	1.5	433	15	AI034859
226	17	1.5	325	134	BB466180	BB466180 BB466180	C 299	17	1.5	433	154	N55392
227	17	1.5	326	134	BB466180	BB466180 BB466180	C 300	17	1.5	433	154	N65550
228	17	1.5	327	134	BB466180	BB466180 BB466180	C 301	17	1.5	434	148	BF417370
229	17	1.5	328	134	BB466180	BB466180 BB466180	C 302	17	1.5	434	210	AQ681704
230	17	1.5	329	134	BB466180	BB466180 BB466180	C 303	17	1.5	435	111	AW147547
231	17	1.5	330	134	BB466180	BB466180 BB466180	C 304	17	1.5	435	122	AW985106
232	17	1.5	331	134	BB466180	BB466180 BB466180	C 305	17	1.5	435	161	BE022226
233	17	1.5	332	134	BB466180	BB466180 BB466180	C 306	17	1.5	436	2	AA142805
234	17	1.5	333	134	BB466180	BB466180 BB466180	C 307	17	1.5	436	17	AI195870
235	17	1.5	334	134	BB466180	BB466180 BB466180	C 308	17	1.5	437	20	AI429361
236	17	1.5	335	134	BB466180	BB466180 BB466180	C 309	17	1.5	437	209	AQ621480
237	17	1.5	336	134	BB466180	BB466180 BB466180	C 310	17	1.5	440	142	BE983034

c 311	17	1.5	441	11	AA730059	n297e03.s	c 384	17	1.5	495	11	AA724665
c 312	17	1.5	442	220	AZ344792	IM0079G06	c 385	17	1.5	495	18	AT255741
c 313	17	1.5	443	4	AA261353	mz86c09.r	c 386	17	1.5	496	15	AI020311
c 314	17	1.5	442	173	BG141818	la94f04.x	c 387	17	1.5	497	105	AL368462
c 315	17	1.5	443	1	AA069268	zf74e05.s	c 388	17	1.5	497	110	AW035003
c 316	17	1.5	443	141	BE954392	UI-M-CC1-	c 389	17	1.5	497	137	BE607656
c 317	17	1.5	444	112	AW185908	se61b05.y	c 390	17	1.5	497	227	B61699
c 318	17	1.5	444	4	AA251107	zs03f06.s	c 391	17	1.5	498	21	AI544917
c 319	17	1.5	446	104	AA1993616	701496796	c 392	17	1.5	499	23	AI641447
c 320	17	1.5	447	9	AA573001	nm33h02.s	c 393	17	1.5	499	164	BE227944
c 321	17	1.5	447	11	AA765232	n277h04.s	c 394	17	1.5	500	104	AI961178
c 322	17	1.5	447	21	AI515812	LD41486.3	c 395	17	1.5	501	105	AL043875
c 323	17	1.5	448	9	AA557540	n102g05.s	c 396	17	1.5	501	169	BF843290
c 324	17	1.5	448	17	AI195217	u162a09.x	c 397	17	1.5	501	216	AZ112010
c 325	17	1.5	449	7	AA418009	zv94g02.x	c 398	17	1.5	501	227	B60752
c 326	17	1.5	451	161	BE019960	sm38a04.y	c 399	17	1.5	502	136	BE515459
c 327	17	1.5	452	153	R7694	H37694	c 400	17	1.5	504	148	BF430257
c 328	17	1.5	452	156	R78075	Y180e04.r1	c 401	17	1.5	504	168	BF735959
c 329	17	1.5	454	115	AA456230	UI-M-BH3-	c 402	17	1.5	504	209	AQ642078
c 330	17	1.5	455	156	T01726	T01726	c 403	17	1.5	506	148	BF442817
c 331	17	1.5	458	2	AA121470	AA121470	c 404	17	1.5	506	217	AZ150238
c 332	17	1.5	458	103	AA860905	w156a04.x	c 405	17	1.5	507	18	AI255846
c 333	17	1.5	458	209	AQ615095	HS-5077.B	c 406	17	1.5	507	158	W54177
c 334	17	1.5	459	171	BF986638	CM0-GN010	c 407	17	1.5	508	137	BE634132
c 335	17	1.5	459	171	BF994822	CM0-GN010	c 408	17	1.5	508	154	N65261
c 336	17	1.5	459	203	AQ221234	HS-2258.B	c 409	17	1.5	508	159	W94720
c 337	17	1.5	459	211	AQ760039	HS-3202.A	c 410	17	1.5	508	206	AQ440700
c 338	17	1.5	459	233	TR204D11Q	AL477125	c 411	17	1.5	508	218	AZ283358
c 339	17	1.5	460	154	N69138	N69138	c 412	17	1.5	510	112	AW210009
c 340	17	1.5	462	115	AAW65549	AW65549	c 413	17	1.5	510	140	BE844948
c 341	17	1.5	464	15	AI060045	UI-R-Cl-k	c 414	17	1.5	514	143	BF063435
c 342	17	1.5	464	17	AI241069	QK05a06.x	c 415	17	1.5	515	20	AI242478
c 343	17	1.5	464	149	BF513352	UI-H-BW1-	c 416	17	1.5	515	215	AZ030987
c 344	17	1.5	467	157	T79571	T79571	c 417	17	1.5	516	103	AI876692
c 345	17	1.5	467	227	B33721	HS-1023-A1-	c 418	17	1.5	516	104	AI994664
c 346	17	1.5	468	14	AA368418	OP49a07.s	c 419	17	1.5	516	118	AW636103
c 347	17	1.5	468	232	FR0016570	AL380283	c 420	17	1.5	516	139	BE749775
c 348	17	1.5	468	232	FR0016570	AL380283	c 421	17	1.5	517	171	AW177395
c 349	17	1.5	469	173	BG096146	BG096146	c 422	17	1.5	517	227	B55730
c 350	17	1.5	470	214	AQ009162	RPCI-23-3	c 423	17	1.5	519	205	AQ335872
c 351	17	1.5	472	139	BE760836	an-0775.A	c 424	17	1.5	520	113	AW317633
c 352	17	1.5	473	4	AA259539	AA259539	c 425	17	1.5	520	164	BE235504
c 353	17	1.5	473	150	BF556164	UI-R-El-f	c 426	17	1.5	521	110	AW066080
c 354	17	1.5	474	104	AI941970	618037D03	c 427	17	1.5	521	149	BF510557
c 355	17	1.5	474	106	AT002537	AT002537	c 428	17	1.5	522	32	AW002924
c 356	17	1.5	476	15	AI032099	AI032099	c 429	17	1.5	524	20	AI464813
c 357	17	1.5	477	1	AA062881	AA062881	c 430	17	1.5	524	149	BF549950
c 358	17	1.5	477	15	AI063668	GH03614.3	c 431	17	1.5	525	18	AI266990
c 359	17	1.5	477	18	AI265270	u197g10.x	c 432	17	1.5	525	24	AI768371
c 360	17	1.5	478	110	AA093668	AA093668	c 433	17	1.5	525	118	AW621631
c 361	17	1.5	478	148	BF459761	BF459761	c 434	17	1.5	526	7	AA450487
c 362	17	1.5	479	17	AI195129	AI195129	c 435	17	1.5	526	22	AI599801
c 363	17	1.5	479	229	AG025795	AG025795	c 436	17	1.5	526	207	AQ531557
c 364	17	1.5	480	17	AI196314	u171a07.y	c 437	17	1.5	527	102	AI812283
c 365	17	1.5	480	20	AI433310	t186d03.x	c 438	17	1.5	527	159	W99065
c 366	17	1.5	480	20	AI433310	t186d03.x	c 439	17	1.5	528	121	AW929923
c 367	17	1.5	480	206	AQ459979	HS-5136.A	c 440	17	1.5	528	150	BF562936
c 368	17	1.5	480	209	AQ659834	Sheared D	c 441	17	1.5	528	211	AQ753770
c 369	17	1.5	480	225	AQ666740	IM0549P03	c 442	17	1.5	528	213	AQ899525
c 370	17	1.5	481	157	T57466	T57466	c 443	17	1.5	529	111	AW113438
c 371	17	1.5	482	7	AA418107	zV94c02.r	c 444	17	1.5	530	118	AW643011
c 372	17	1.5	483	15	AI012493	EST206944	c 445	17	1.5	530	207	AQ515614
c 373	17	1.5	484	17	AI194995	u159c07.x	c 446	17	1.5	530	213	AQ944572
c 374	17	1.5	485	16	AI159449	v279c12.r	c 447	17	1.5	532	12	AA798989
c 375	17	1.5	486	210	AQ0717243	HS-2152.B	c 448	17	1.5	532	208	AQ588331
c 376	17	1.5	489	1	AA028595	ml13d01.r	c 449	17	1.5	533	16	AI120993
c 377	17	1.5	491	110	AW054491	660008A09	c 450	17	1.5	533	205	AZ073019
c 378	17	1.5	491	102	AI787255	uJ58b05.x	c 451	17	1.5	534	30	AV527110
c 379	17	1.5	492	105	AJ397095	AJ397095	c 452	17	1.5	534	111	AW113436
c 380	17	1.5	493	104	AI960412	sc8e02.y	c 453	17	1.5	534	240	AQ706476
c 381	17	1.5	493	154	N75182	Y262g12.r1	c 454	17	1.5	535	147	BF357340
c 382	17	1.5	493	215	AZ076812	RPCI-23-4	c 455	17	1.5	535	158	W44267
c 383	17	1.5	494	6	AA397312	mr39e10.r	c 456	17	1.5	536	110	AW092056

c 457	17	1.5	536	205	AQ373309	RPC111-14	530	17	1.5	592	172	BG030977	BG030977	602299110
c 458	17	1.5	536	209	AQ662468	HS_5455_A	c 531	17	1.5	593	122	AW974223	AW974223	EST386326
c 459	17	1.5	537	104	A1993303	701496008	532	17	1.5	594	164	BE204877	BE204877	EST397617
c 460	17	1.5	538	14	AA968386	uc71h04.x	533	17	1.5	595	121	A2416957	A2416957	1M0192K21
c 461	17	1.5	538	167	BE464128	hy21b02.x	534	17	1.5	595	121	BE025420	BE025420	945026H05
c 462	17	1.5	538	211	AQ0808101	HS_2169_B	c 535	17	1.5	598	115	AW444001	AW444001	EST308931
c 463	17	1.5	539	10	AA647314	vs91b09.r	c 536	17	1.5	599	210	AQ0706542	AQ0706542	HS_5534_B
c 464	17	1.5	540	18	A1265719	uj05e06.x	c 537	17	1.5	600	173	BG097735	BG097735	EST462254
c 465	17	1.5	540	229	AQ025569	Oryza sat	c 538	17	1.5	601	122	AW944731	AW944731	EST336781
c 466	17	1.5	540	231	CNS03003	Tetraodon	c 539	17	1.5	601	213	AQ923704	AQ923704	RPCI-23-2
c 467	17	1.5	542	15	A1042807	uc75e02.x	540	17	1.5	602	10	AA660259	AA660259	00128 Mtr
c 468	17	1.5	542	18	A1265477	uj02e04.x	541	17	1.5	602	15	AI035891	AI035891	ud48h10.r
c 469	17	1.5	544	110	AW074099	xb07g02.x	c 542	17	1.5	604	118	AW638058	AW638058	bl65f10.w
c 470	17	1.5	544	111	AW163107	au97d09.x	543	17	1.5	604	205	AQ385604	AQ385604	RPC111-14
c 471	17	1.5	545	172	BG020053	dc70g08.x	544	17	1.5	604	225	AQ2664964	AQ2664964	1M0545H10
c 472	17	1.5	545	213	AQ089182	HS_3214_A	545	17	1.5	605	171	AZ185411	AZ185411	SP_1005_A
c 473	17	1.5	546	110	AW096807	EST289987	546	17	1.5	606	223	AZ575490	AZ575490	HSC_00232
c 474	17	1.5	546	115	AW413067	uq51g10.x	547	17	1.5	607	205	AQ2323192	AQ2323192	RPC111-11
c 475	17	1.5	546	140	BE821887	GMT00015B	c 548	17	1.5	609	16	AI132255	AI132255	ue32h03.x
c 476	17	1.5	547	137	BE634152	uv69f01.x	c 549	17	1.5	610	117	AW561958	AW561958	T8221 MVA
c 477	17	1.5	548	167	BE436637	EST407755	550	17	1.5	611	8	AA544037	AA544037	vk26g06.r
c 478	17	1.5	549	143	BF064099	7h94a04.x	c 551	17	1.5	611	18	AI256575	AI256575	ui88b06.x
c 479	17	1.5	549	165	BE288257	601095435	552	17	1.5	611	173	BG136066	BG136066	EST476508
c 480	17	1.5	549	216	AZ105006	RPCI-23-4	553	17	1.5	612	226	BQ3675	BQ3675	CSRL-184G1
c 481	17	1.5	549	233	TA145C11P	T. brucei	554	17	1.5	614	210	AQ702053	AQ702053	HS_5400_B
c 482	17	1.5	550	115	AW408064	UI-HF-BM0	c 555	17	1.5	615	110	AW067187	AW067187	683017B03
c 483	17	1.5	550	217	AZ216062	Sheared D	c 556	17	1.5	615	118	AW637911	AW637911	bl64a02.w
c 484	17	1.5	551	142	BE997918	EST429641	c 557	17	1.5	616	220	AZ363679	AZ363679	1M0109L14
c 485	17	1.5	551	154	N79232	N79232 za12h09.r1	558	17	1.5	617	18	AI256169	AI256169	ui95g10.x
c 486	17	1.5	552	114	AW399173	EST309673	c 559	17	1.5	617	205	AQ383054	AQ383054	RPC111-13
c 487	17	1.5	552	233	TA272H05Q	AL487849 T. brucei	c 560	17	1.5	618	17	AI172961	AI172961	ud62g12.x
c 488	17	1.5	553	1	AA050383	AA050383 mj11h04.r	c 561	17	1.5	618	17	AI195139	AI195139	ui61b05.x
c 489	17	1.5	554	1	AA009280	mp96c01.r	c 562	17	1.5	619	18	AI255443	AI255443	ui85f02.x
c 490	17	1.5	554	203	AQ238601	RPC111-70	c 563	17	1.5	619	118	AW685923	AW685923	NF036604N
c 491	17	1.5	556	115	AW413011	uq49g11.x	564	17	1.5	619	214	AQ959509	AQ959509	LRELO03TR
c 492	17	1.5	556	218	AZ247106	RPCI-23-4	565	17	1.5	620	2	AA118625	AA118625	mp58g07.r
c 493	17	1.5	557	23	A1652234	wb28a07.x	566	17	1.5	620	30	AVS20694	AVS20694	AVS20694
c 494	17	1.5	557	222	AZ478076	1M0297120	c 567	17	1.5	620	149	BF531607	BF531607	602091272
c 495	17	1.5	558	114	AW606014	660032G05	c 568	17	1.5	621	16	AI114886	AI114886	ui49e10.x
c 496	17	1.5	561	159	W98929	mf91c09.r1	569	17	1.5	623	216	AZ093561	AZ093561	RPCI-23-4
c 497	17	1.5	562	112	AZ230742	u067c12.y	c 570	17	1.5	623	224	AZ594165	AZ594165	1M0406C17
c 498	17	1.5	562	173	BG134198	EST467090	571	17	1.5	626	118	AW651499	AW651499	EST329953
c 499	17	1.5	565	118	AW626225	EST320132	c 572	17	1.5	627	206	AQ460117	AQ460117	HS_51233_B
c 500	17	1.5	566	17	A1218888	qg82c04.x	573	17	1.5	628	111	AW116720	AW116720	fi19a11.x
c 501	17	1.5	567	18	A1265721	uj05e08.x	574	17	1.5	629	2	AA122844	AA122844	mq05d05.r
c 502	17	1.5	569	121	AW932478	EST358321	c 575	17	1.5	629	120	AW806505	AW806505	ILO-ST000
c 503	17	1.5	570	225	AZ644517	1M0508D02	576	17	1.5	630	110	AW076310	AW076310	614065D09
c 504	17	1.5	571	145	BF232367	de20c02.x	577	17	1.5	630	215	AZ054681	AZ054681	RPCI-23-4
c 505	17	1.5	571	171	BF952845	MR3-NNO21	c 578	17	1.5	631	18	AI266789	AI266789	uj07b02.x
c 506	17	1.5	572	219	AZ314062	1M0030L12	c 579	17	1.5	631	150	BF612123	BF612123	G891d08.y
c 507	17	1.5	572	3	AA207945	mv92g01.r	c 580	17	1.5	631	209	AQ613807	AQ613807	HS_5118_B
c 508	17	1.5	572	173	BG078507	H3028E04 -	c 581	17	1.5	631	216	AZ106729	AZ106729	RPCI-23-4
c 509	17	1.5	572	12	AA794556	vu68h09.r	c 582	17	1.5	632	103	AI892189	AI892189	wt188b01.y
c 510	17	1.5	573	167	BE472037	EST416890	c 583	17	1.5	632	221	AZ440819	AZ440819	1M0232N06
c 511	17	1.5	575	142	BF019570	ux12g02.y	c 584	17	1.5	633	165	BE310403	BE310403	601088049
c 512	17	1.5	577	32	AW017996	614065D09	c 585	17	1.5	636	111	AW120212	AW120212	614086F12
c 513	17	1.5	577	115	AW453387	660032G05	c 586	17	1.5	639	121	AW930794	AW930794	EST356637
c 514	17	1.5	577	210	AQ0707925	HS_5561_B	c 587	17	1.5	640	21	AI515520	AI515520	LD47387.5
c 515	17	1.5	580	143	BF065704	CEB001	c 588	17	1.5	640	213	AQ933323	AQ933323	RPCI-23-2
c 516	17	1.5	581	122	AW971521	EST383610	c 589	17	1.5	642	208	AQ559628	AQ559628	HS_5222_B
c 517	17	1.5	581	140	BE868551	601444584	c 590	17	1.5	644	147	BF384594	BF384594	602046863
c 518	17	1.5	581	208	AQ542743	RPCI-11-3	c 591	17	1.5	646	224	AZ608864	AZ608864	1M0433P18
c 519	17	1.5	581	208	AQ583722	RPCI-11-4	c 592	17	1.5	647	173	BG124887	BG124887	EST470533
c 520	17	1.5	582	115	AW429221	EST306677	c 593	17	1.5	648	214	AQ959508	AQ959508	LELE03TF
c 521	17	1.5	582	105	AL501131	AL501131	594	17	1.5	654	103	AI892095	AI892095	wt11h04.y
c 522	17	1.5	586	110	AW055114	wy99g12.x	595	17	1.5	654	165	BE306066	BE306066	601101772
c 523	17	1.5	586	112	AW217632	EST296346	596	17	1.5	656	31	AV699874	AV699874	AV699874
c 524	17	1.5	587	4	AA221344	mw11e06.r	597	17	1.5	657	31	AV700783	AV700783	AV700783
c 525	17	1.5	587	173	BG127124	EST472770	598	17	1.5	659	171	BF983209	BF983209	602308235
c 526	17	1.5	589	173	BG127124	EST472770	599	17	1.5	665	221	AZ403388	AZ403388	1M0171F10
c 527	17	1.5	591	119	AW04974	sk56h09.y	600	17	1.5	667	122	AW955017	AW955017	EST367087
c 528	17	1.5	592	104	AI976220	XC28b12.x	601	17	1.5	667	145	BF223540	BF223540	7q77d04.x
c 529	17	1.5	592	110	AW080944	xc28b12.x	c 602	17	1.5	668	122	AW977244	AW977244	EST389353

c 603	17	1.5	668	148	BF472550	UI-M-BH3-	676	17	1.5	867	149	BF540272	602052326
c 604	17	1.5	669	151	AI055307	coau0003K	677	17	1.5	869	231	CNS03YLD	AL266458 Tetraodon
c 605	17	1.5	673	117	AW561932	T8190 MVA	c 678	17	1.5	874	151	BF673826	602135976
c 606	17	1.5	678	117	AW561940	T8199 MVA	c 679	17	1.5	881	151	BF678162	602084828
c 607	17	1.5	679	32	AW760013	AV760013	c 680	17	1.5	883	144	BF135191	601780248
c 608	17	1.5	684	106	AU004136	AU004136	c 681	17	1.5	885	223	AZ539819	ENTDI47TR
c 609	17	1.5	684	219	AQ965701	LERID81TF	c 682	17	1.5	886	222	CNS013LO	AI103014 Drosophila
c 610	17	1.5	688	219	AZ330625	LM0056N05	c 683	17	1.5	897	122	AW983240	HVSN09000
c 611	17	1.5	689	218	AZ281740	RPCI-23-1	c 684	17	1.5	904	144	BF164918	BF164918
c 612	17	1.5	690	29	AV405031	AV405031	c 685	17	1.5	905	145	BF207807	601862510
c 613	17	1.5	690	105	AL157608	DKF2P761J	c 686	17	1.5	908	231	CNS040A6	AL299751 Tetraodon
c 614	17	1.5	693	20	AI140485	GH24226.5	c 687	17	1.5	919	150	BF578341	602092705
c 615	17	1.5	693	159	W85447	W85447 mf45d12.r1	c 688	17	1.5	920	173	BF109858	602279524
c 616	17	1.5	693	205	AQ368186	AQ368186	c 689	17	1.5	923	230	CNS02S9D	AL211594 Tetraodon
c 617	17	1.5	693	218	AZ280318	RPCI-23-1	c 690	17	1.5	928	217	AZ207403	SP_0135_A
c 618	17	1.5	700	105	AL507349	AL507349	c 691	17	1.5	930	105	AL041339	AL041339 DKF2P434P
c 619	17	1.5	703	117	AW561945	AW561945	c 692	17	1.5	932	12	AA826243	off1b07.s
c 620	17	1.5	703	207	AQ330613	AQ330613	c 693	17	1.5	932	145	BF206687	601871052
c 621	17	1.5	704	139	BE789361	BE789361	c 694	17	1.5	932	232	CNS05CW9	AL331650 Tetraodon
c 622	17	1.5	705	16	AI147198	AI147198 q89b02.x	c 695	17	1.5	947	174	BG180262	602329880
c 623	17	1.5	706	162	BE039368	BE039368	c 696	17	1.5	951	230	CNS03BIB	AL236540 Tetraodon
c 624	17	1.5	708	219	AZ321391	AI0041E02	c 697	17	1.5	953	144	BF158834	601766185
c 625	17	1.5	711	205	AQ387022	AQ387022	c 698	17	1.5	965	232	CNS0581Q	AL325979 Tetraodon
c 626	17	1.5	715	105	AJ393511	AJ393511	c 699	17	1.5	966	231	CNS046Y2	AL277283 Tetraodon
c 627	17	1.5	716	221	AZ422750	IM0201N11	c 700	17	1.5	969	172	BG025210	602276308
c 628	17	1.5	717	214	AQ960050	LERFB10TR	c 701	17	1.5	982	166	BE374662	601224161
c 629	17	1.5	719	229	AG022764	Oryza sat	c 702	17	1.5	986	136	BE536779	601065666
c 630	17	1.5	720	149	BF527109	BF527109	c 703	17	1.5	996	136	BE540409	601065581
c 631	17	1.5	720	229	CNS00VPD	AL092575 Arabidops	c 704	17	1.5	996	143	BF028596	601763819
c 632	17	1.5	726	104	AI968212	AI968212 wu14d12.x	c 705	17	1.5	1002	139	BE746027	601578552
c 633	17	1.5	730	219	AZ298879	AZ298879	c 706	17	1.5	1003	151	BF676115	602084104
c 634	17	1.5	732	112	AW243107	AW243107	c 707	17	1.5	1015	231	CNS03R2S	AL256717 Tetraodon
c 635	17	1.5	735	144	BF144955	601790248	c 708	17	1.5	1021	140	BE871658	601449502
c 636	17	1.5	737	214	AQ960049	AQ960049	c 709	17	1.5	1022	230	CNS0367T	AL229682 Tetraodon
c 637	17	1.5	743	205	AQ389624	AQ389624	c 710	17	1.5	1025	174	EG245523	602355082
c 638	17	1.5	746	170	BR860061	BR860061	c 711	17	1.5	1033	172	BG026775	602294239
c 639	17	1.5	747	102	AI806541	AI806541	c 712	17	1.5	1033	231	CNS049S2	AL280955 Tetraodon
c 640	17	1.5	753	170	BF860057	BF860057	c 713	17	1.5	1044	230	CNS027K2	AL184764 Tetraodon
c 641	17	1.5	757	211	AQ749736	HS_5573_A	c 714	17	1.5	1050	150	BF580202	602099016
c 642	17	1.5	757	220	AZ381609	AZ381609	c 715	17	1.5	1067	232	CNS05821	AL326566 Tetraodon
c 643	17	1.5	763	23	AI552079	AI552079 wD27f11.x	c 716	17	1.5	1077	232	CNS042RX	AL314646 Tetraodon
c 644	17	1.5	763	146	BF265754	BF265754 HV_Cea001	c 717	17	1.5	1084	142	BE959146	601645234
c 645	17	1.5	765	217	AI185693	AI185693 SP_1005_B	c 718	17	1.5	1088	232	CNS05SBB	AL331638 Tetraodon
c 646	17	1.5	765	16	AI098594	AI098594 ue31e07.y	c 719	17	1.5	1100	232	CNS05FDB	AL334849 Tetraodon
c 647	17	1.5	767	203	AQ245753	AQ245753 HS_2060_B	c 720	17	1.5	1295	151	BF689196	602184871
c 648	17	1.5	768	140	BE856574	BE856574	c 721	17	1.5	1295	151	BF689196	602184871
c 649	17	1.5	770	165	BE278980	BE278980	c 722	17	1.5	1572	176	BF310358	601895193
c 650	17	1.5	771	103	AI892884	AI892884	c 723	17	1.5	1706	136	BE538441	601068280
c 651	17	1.5	774	14	AI005926	AI005926 ua84d03.y	c 724	17	1.5	1756	144	BF129804	601811263
c 652	17	1.5	774	106	AL534495	AL534495	c 725	17	1.5	2542	166	BE420611	HWM000.F1
c 653	17	1.5	774	144	BE165359	BE165359	c 726	16	1.4	46	8	AA485011	aa39e10.r
c 654	17	1.5	785	137	BE613963	BE613963	c 727	16	1.4	96	224	AZ581528	IM0370824
c 655	17	1.5	787	171	BF977226	BF977226	c 728	16	1.4	110	12	AA792585	AA792585 vs86e05.r
c 656	17	1.5	788	173	BG121376	BG121376	c 729	16	1.4	113	21	AI551092	vx36g12.x
c 657	17	1.5	789	102	AI831616	w139q11.x	c 730	16	1.4	113	139	BE772861	RC1-FT013
c 658	17	1.5	795	205	AQ364209	AQ364209 nbxb0060L	c 731	16	1.4	116	139	BE772863	RC1-FT013
c 659	17	1.5	799	206	AQ397015	AQ397015	c 732	16	1.4	121	7	AA450380	zx05a06.r
c 660	17	1.5	803	107	AI013198	AI013198	c 733	16	1.4	127	114	AW342814	f189a02.x
c 661	17	1.5	809	162	BE039271	BE039271 AC01D02_A	c 734	16	1.4	132	219	AZ287647	RPCI-23-1
c 662	17	1.5	810	15	AI031673	AI031673	c 735	16	1.4	134	162	BE086616	QV1-BT067
c 663	17	1.5	810	142	BE971489	BE971489	c 736	16	1.4	137	4	AA247125	cop1025.s
c 664	17	1.5	810	149	BF530598	BF530598	c 737	16	1.4	137	108	AV079404	AV079404
c 665	17	1.5	812	107	AI049591	AI049591	c 738	16	1.4	137	142	BE993222	UI-M-B21-
c 666	17	1.5	816	138	BE727424	BE727424	c 739	16	1.4	139	147	BF380183	MR3-UT005
c 667	17	1.5	818	171	BF964828	BF964828	c 740	16	1.4	141	139	BE772862	RC1-FT013
c 668	17	1.5	825	230	CNS036JM	AL230107 Tetraodon	c 741	16	1.4	143	147	BF356364	QV4-HT089
c 669	17	1.5	825	231	CNS04FXV	AL288940 Tetraodon	c 742	16	1.4	146	2	AA144030	ms14e09.r
c 670	17	1.5	844	217	AZ197427	AZ197427	c 743	16	1.4	148	147	BF356331	QV4-HT089
c 671	17	1.5	846	217	AI187672	AI187672	c 744	16	1.4	149	218	AZ276748	RPCI-23-1
c 672	17	1.5	848	102	AI790802	AI790802	c 745	16	1.4	150	139	BE772860	RC1-FT013
c 673	17	1.5	852	21	AI528304	AI528304	c 746	16	1.4	151	147	BF387485	UI-R-CA1-
c 674	17	1.5	855	223	AZ535642	AZ535642	c 747	16	1.4	156	16	AI145467	UI-R-BT0-
c 675	17	1.5	867	144	BF128895	BF128895	c 748	16	1.4	156	147	BF356326	QV4-HT089

c 749	16	1.4	160	218	AZ238159	822	16	1.4	241	152	H05810
c 750	16	1.4	161	4	AA248662	c 823	16	1.4	242	13	AA926891
c 751	16	1.4	165	224	AZ580176	824	16	1.4	242	130	BB333314
c 752	16	1.4	169	232	FR0025349	825	16	1.4	242	170	BF889272
c 753	16	1.4	171	146	BF323579	c 826	16	1.4	243	11	AA767516
c 754	16	1.4	172	132	BB398539	827	16	1.4	243	28	AV337181
c 755	16	1.4	173	4	AZ370559	c 828	16	1.4	243	119	AW03124
c 756	16	1.4	175	223	AZ576543	829	16	1.4	244	152	D63043
c 757	16	1.4	180	175	C41259	830	16	1.4	244	152	D63043
c 758	16	1.4	182	121	AW929818	831	16	1.4	245	152	D79784
c 759	16	1.4	182	143	BF087233	832	16	1.4	245	152	D79784
c 760	16	1.4	183	30	AV118027	833	16	1.4	245	152	D79784
c 761	16	1.4	184	14	AA981817	834	16	1.4	245	152	D79784
c 762	16	1.4	184	164	BE265899	c 835	16	1.4	246	30	AV527529
c 763	16	1.4	188	170	BF901736	836	16	1.4	246	124	BB060689
c 764	16	1.4	189	219	AZ322859	837	16	1.4	247	10	AA686272
c 765	16	1.4	191	109	AV140088	c 838	16	1.4	248	26	AV523623
c 766	16	1.4	192	1	AA025026	839	16	1.4	248	149	BF543046
c 767	16	1.4	193	29	AV353922	840	16	1.4	248	227	B76912
c 768	16	1.4	193	121	AW868228	c 841	16	1.4	249	118	AW829858
c 769	16	1.4	193	170	BF872638	c 842	16	1.4	249	156	R75485
c 770	16	1.4	195	152	F04911	843	16	1.4	250	111	AW120235
c 771	16	1.4	196	30	AV533359	844	16	1.4	250	116	AW494505
c 772	16	1.4	196	30	AV536890	845	16	1.4	250	116	AW494505
c 773	16	1.4	196	158	W17868	c 846	16	1.4	250	122	AW955526
c 774	16	1.4	197	158	W70428	c 847	16	1.4	250	134	BB467857
c 775	16	1.4	199	14	A1004749	c 848	16	1.4	251	11	AA742005
c 776	16	1.4	199	139	BE745018	849	16	1.4	252	227	B62360
c 777	16	1.4	200	139	BE745018	850	16	1.4	252	106	AV053840
c 778	16	1.4	202	207	AQ479394	851	16	1.4	253	212	AQ849662
c 779	16	1.4	203	3	AA210130	852	16	1.4	253	214	AQ849662
c 780	16	1.4	203	28	AV326051	c 853	16	1.4	254	214	AQ849662
c 781	16	1.4	203	171	BF930644	c 854	16	1.4	255	146	BF332951
c 782	16	1.4	205	127	BB205876	c 855	16	1.4	256	23	AI661255
c 783	16	1.4	205	170	BF879927	856	16	1.4	257	18	AI277153
c 784	16	1.4	209	32	AW030179	c 857	16	1.4	258	7	AA404326
c 785	16	1.4	209	126	BB152227	858	16	1.4	259	152	D61842
c 786	16	1.4	211	24	A1720037	c 859	16	1.4	260	11	AA721294
c 787	16	1.4	212	170	BF910283	c 860	16	1.4	261	157	T94881
c 788	16	1.4	214	114	AA368408	861	16	1.4	262	26	AV240713
c 789	16	1.4	216	160	BB583977	862	16	1.4	262	107	AV012360
c 790	16	1.4	217	170	BF889266	863	16	1.4	262	109	AV143815
c 791	16	1.4	222	AZ475577	864	16	1.4	263	157	T76059	
c 792	16	1.4	219	150	BF599807	c 865	16	1.4	263	230	CNS02MVP
c 793	16	1.4	221	6	AA367181	c 866	16	1.4	264	29	AV378837
c 794	16	1.4	221	26	AV239765	867	16	1.4	264	109	AV119132
c 795	16	1.4	221	28	AV313071	c 868	16	1.4	265	14	AA987824
c 796	16	1.4	222	154	N39861	c 869	16	1.4	266	30	AV442139
c 797	16	1.4	222	171	BF988822	c 870	16	1.4	266	6	AA368824
c 798	16	1.4	223	204	AQ283439	c 871	16	1.4	267	8	AA505143
c 799	16	1.4	224	114	AW351715	c 872	16	1.4	267	28	AV313238
c 800	16	1.4	224	161	BB598512	873	16	1.4	268	109	AV091973
c 801	16	1.4	225	19	AI380545	c 874	16	1.4	268	169	BF812215
c 802	16	1.4	226	18	A1265981	875	16	1.4	269	8	AA504435
c 803	16	1.4	228	164	BE206851	c 876	16	1.4	269	114	AW368407
c 804	16	1.4	229	17	AI165232	c 877	16	1.4	269	131	BB354151
c 805	16	1.4	230	109	AV148104	c 878	16	1.4	270	114	AW368395
c 806	16	1.4	232	115	AW477867	879	16	1.4	271	26	AV225168
c 807	16	1.4	233	3	AA172826	880	16	1.4	271	28	AV340570
c 808	16	1.4	234	116	AW522364	881	16	1.4	271	108	AV062295
c 809	16	1.4	235	30	AV534313	882	16	1.4	271	108	AV081165
c 810	16	1.4	235	134	BB498656	883	16	1.4	271	209	AO659163
c 811	16	1.4	236	124	BB075799	884	16	1.4	272	170	BF895458
c 812	16	1.4	237	26	AV250410	c 885	16	1.4	272	227	B62300
c 813	16	1.4	237	28	AV321965	c 886	16	1.4	273	109	AV125961
c 814	16	1.4	237	154	N53196	c 887	16	1.4	273	116	AW485055
c 815	16	1.4	238	148	BF458313	c 888	16	1.4	273	163	BE160960
c 816	16	1.4	239	26	AV248977	889	16	1.4	273	173	BG131313
c 817	16	1.4	240	124	BB059554	c 890	16	1.4	274	135	BB537923
c 818	16	1.4	240	3	AA195501	891	16	1.4	274	148	BF452306
c 819	16	1.4	240	116	AW499673	892	16	1.4	275	108	AV074408
c 820	16	1.4	241	20	AI410144	893	16	1.4	276	131	BB342248
c 821	16	1.4	241	144	BF149259	894	16	1.4	277	28	AV328685

c 895	16	1.4	277	112	AW219422	EST302000
c 896	16	1.4	277	120	AW818446	RC1-ST027
c 897	16	1.4	277	126	BB166656	BB166656
c 898	16	1.4	277	130	TG32383	TgEST0147 T
c 899	16	1.4	278	137	BB3222704	BB3222704
c 900	16	1.4	278	149	BF545581	U1-R-CI-J
c 901	16	1.4	278	152	D62789	HOM325H1B
c 902	16	1.4	278	170	BF855858	PW2-FN021
c 903	16	1.4	279	201	AQ0117475	CIT-HSP-2
c 904	16	1.4	279	108	AV031765	AV031765
c 905	16	1.4	279	123	BB037926	BB037926
c 906	16	1.4	279	170	BF900485	PM1-WT019
c 907	16	1.4	281	5	AA300006	EST12694
c 908	16	1.4	282	106	AU069811	AU069811
c 909	16	1.4	282	107	AU164359	AU164359
c 910	16	1.4	282	108	AV0115535	AV0115535
c 911	16	1.4	282	123	BB040834	BB040834
c 912	16	1.4	282	171	BF964496	RC1-NN023
c 913	16	1.4	282	171	BF964520	RC1-NN023
c 914	16	1.4	283	109	AV089625	AV089625
c 915	16	1.4	283	109	AV125777	AV125777
c 916	16	1.4	283	115	AW435651	76860 MAR
c 917	16	1.4	283	122	BB013380	BB013380
c 918	16	1.4	283	127	BB178149	BB178149
c 919	16	1.4	283	131	BB346268	BB346268
c 920	16	1.4	283	138	BE79662	dF51f04.Y
c 921	16	1.4	284	146	BF113451	601900042
c 922	16	1.4	284	147	BF359787	RCO-MT005
c 923	16	1.4	284	226	AZ724728	RPC1-24-1
c 924	16	1.4	285	10	AA685645	EST107735
c 925	16	1.4	285	124	BB092266	BB092266
c 926	16	1.4	286	124	BB060210	BB060210
c 927	16	1.4	286	125	BB115535	BB115535
c 928	16	1.4	286	128	BB237597	BB237597
c 929	16	1.4	286	128	BB255072	BB255072
c 930	16	1.4	286	130	BB305124	BB305124
c 931	16	1.4	286	135	BB521513	BB521513
c 932	16	1.4	287	109	AV147056	AV147056
c 933	16	1.4	287	122	BB007100	BB007100
c 934	16	1.4	287	128	BB244664	BB244664
c 935	16	1.4	287	131	BB525916	BB525916
c 936	16	1.4	287	134	BB483501	BB483501
c 937	16	1.4	288	109	AV136580	AV136580
c 938	16	1.4	289	32	AV740252	AV740252
c 939	16	1.4	289	129	BB268596	BB268596
c 940	16	1.4	289	204	AQ310966	CIT-HSP-2
c 941	16	1.4	290	25	AV156408	AV156408
c 942	16	1.4	290	106	AT002357	AT002357
c 943	16	1.4	290	106	AU075716	AU075716
c 944	16	1.4	290	109	AV117651	AV117651
c 945	16	1.4	290	124	BB068245	BB068245
c 946	16	1.4	291	120	AW792869	CM3-UW000
c 947	16	1.4	291	125	BB130461	BB130461
c 948	16	1.4	291	130	BB308245	BB308245
c 949	16	1.4	291	147	BF359786	RCO-MT005
c 950	16	1.4	292	132	BB395107	BB395107
c 951	16	1.4	292	171	BF986988	RC1-GN002
c 952	16	1.4	292	220	AZ383038	RC1-L0L20
c 953	16	1.4	292	227	BB7364	BB7364
c 954	16	1.4	293	3	AA188769	zp78g08.r
c 955	16	1.4	293	16	AA103420	EST212709
c 956	16	1.4	293	110	AW043123	SW29F05 P
c 957	16	1.4	293	170	BF989531	PM1-WT018
c 958	16	1.4	294	130	BB322638	BB322638
c 959	16	1.4	294	160	BB570181	BB570181
c 960	16	1.4	294	170	BF855886	PM3-TN010
c 961	16	1.4	295	25	AV165130	AV165130
c 962	16	1.4	296	12	AA804368	nz30a12.s
c 963	16	1.4	296	25	AV153751	AV153751
c 964	16	1.4	296	27	AV302868	AV302868
c 965	16	1.4	296	28	AV337980	AV337980
c 966	16	1.4	296	108	AV075372	AV075372
c 967	16	1.4	296	124	BB079428	BB079428
c 895	16	1.4	277	112	AW219422	EST302000
c 896	16	1.4	277	120	AW818446	RC1-ST027
c 897	16	1.4	277	126	BB166656	BB166656
c 898	16	1.4	277	130	TG32383	TgEST0147 T
c 899	16	1.4	278	137	BB3222704	BB3222704
c 900	16	1.4	278	149	BF545581	U1-R-CI-J
c 901	16	1.4	278	152	D62789	HOM325H1B
c 902	16	1.4	278	170	BF855858	PW2-FN021
c 903	16	1.4	279	201	AQ0117475	CIT-HSP-2
c 904	16	1.4	279	108	AV031765	AV031765
c 905	16	1.4	279	123	BB037926	BB037926
c 906	16	1.4	279	170	BF900485	PM1-WT019
c 907	16	1.4	281	5	AA300006	EST12694
c 908	16	1.4	282	106	AU069811	AU069811
c 909	16	1.4	282	107	AU164359	AU164359
c 910	16	1.4	282	108	AV0115535	AV0115535
c 911	16	1.4	282	123	BB040834	BB040834
c 912	16	1.4	282	171	BF964496	RC1-NN023
c 913	16	1.4	282	171	BF964520	RC1-NN023
c 914	16	1.4	283	109	AV089625	AV089625
c 915	16	1.4	283	109	AV125777	AV125777
c 916	16	1.4	283	115	AW435651	76860 MAR
c 917	16	1.4	283	122	BB013380	BB013380
c 918	16	1.4	283	127	BB178149	BB178149
c 919	16	1.4	283	131	BB346268	BB346268
c 920	16	1.4	283	138	BE79662	dF51f04.Y
c 921	16	1.4	284	146	BF113451	601900042
c 922	16	1.4	284	147	BF359787	RCO-MT005
c 923	16	1.4	284	226	AZ724728	RPC1-24-1
c 924	16	1.4	285	10	AA685645	EST107735
c 925	16	1.4	285	124	BB092266	BB092266
c 926	16	1.4	286	124	BB060210	BB060210
c 927	16	1.4	286	125	BB115535	BB115535
c 928	16	1.4	286	128	BB237597	BB237597
c 929	16	1.4	286	128	BB255072	BB255072
c 930	16	1.4	286	130	BB305124	BB305124
c 931	16	1.4	286	135	BB521513	BB521513
c 932	16	1.4	287	109	AV147056	AV147056
c 933	16	1.4	287	122	BB007100	BB007100
c 934	16	1.4	287	128	BB244664	BB244664
c 935	16	1.4	287	131	BB525916	BB525916
c 936	16	1.4	287	134	BB483501	BB483501
c 937	16	1.4	288	109	AV136580	AV136580
c 938	16	1.4	289	32	AV740252	AV740252
c 939	16	1.4	289	129	BB268596	BB268596
c 940	16	1.4	289	204	AQ310966	CIT-HSP-2
c 941	16	1.4	290	25	AV156408	AV156408
c 942	16	1.4	290	106	AT002357	AT002357
c 943	16	1.4	290	106	AU075716	AU075716
c 944	16	1.4	290	109	AV117651	AV117651
c 945	16	1.4	290	124	BB068245	BB068245
c 946	16	1.4	291	120	AW792869	CM3-UW000
c 947	16	1.4	291	125	BB130461	BB130461
c 948	16	1.4	291	130	BB308245	BB308245
c 949	16	1.4	291	147	BF359786	RCO-MT005
c 950	16	1.4	292	132	BB395107	BB395107
c 951	16	1.4	292	171	BF986988	RC1-GN002
c 952	16	1.4	292	220	AZ383038	RC1-L0L20
c 953	16	1.4	292	227	BB7364	BB7364
c 954	16	1.4	293	3	AA188769	zp78g08.r
c 955	16	1.4	293	16	AA103420	EST212709
c 956	16	1.4	293	110	AW043123	SW29F05 P
c 957	16	1.4	293	170	BF989531	PM1-WT018
c 958	16	1.4	294	130	BB322638	BB322638
c 959	16	1.4	294	160	BB570181	BB570181
c 960	16	1.4	294	170	BF855886	PM3-TN010
c 961	16	1.4	295	25	AV165130	AV165130
c 962	16	1.4	296	12	AA804368	nz30a12.s
c 963	16	1.4	296	25	AV153751	AV153751
c 964	16	1.4	296	27	AV302868	AV302868
c 965	16	1.4	296	28	AV337980	AV337980
c 966	16	1.4	296	108	AV075372	AV075372
c 967	16	1.4	296	124	BB079428	BB079428
c 895	16	1.4	277	112	AW219422	EST302000
c 896	16	1.4	277	120	AW818446	RC1-ST027
c 897	16	1.4	277	126	BB166656	BB166656
c 898	16	1.4	277	130	TG32383	TgEST0147 T
c 899	16	1.4	278	137	BB3222704	BB3222704
c 900	16	1.4	278	149	BF545581	U1-R-CI-J
c 901	16	1.4	278	152	D62789	HOM325H1B
c 902	16	1.4	278	170	BF855858	PW2-FN021
c 903	16	1.4	279	201	AQ0117475	CIT-HSP-2
c 904	16	1.4	279	108	AV031765	AV031765
c 905	16	1.4	279	123	BB037926	BB037926
c 906	16	1.4	279	170	BF900485	PM1-WT019
c 907	16	1.4	281	5	AA300006	EST12694
c 908	16	1.4	282	106	AU069811	AU069811
c 909	16	1.4	282	107	AU164359	AU164359
c 910	16	1.4	282	108	AV0115535	AV0115535
c 911	16	1.4	282	123	BB040834	BB040834
c 912	16	1.4	282	171	BF964496	RC1-NN023
c 913	16	1.4	282	171	BF964520	RC1-NN023
c 914	16	1.4	283	109	AV089625	AV089625
c 915	16	1.4	283	109	AV125777	AV125777
c 916	16	1.4	283	115	AW435651	76860 MAR
c 917	16	1.4	283	122	BB013380	BB013380
c 918	16	1.4	283	127	BB178149	BB178149
c 919	16	1.4	283	131	BB346268	BB346268
c 920	16	1.4	283	138	BE79662	dF51f04.Y
c 921	16	1.4	284	146	BF113451	601900042
c 922	16	1.4	284	147	BF359787	RCO-MT005
c 923	16	1.4	284	226	AZ724728	RPC1-24-1
c 924	16	1.4	285	10	AA685645	EST107735
c 925	16	1.4	285	124	BB092266	BB092266
c 926	16	1.4	286	124	BB060210	BB060210
c 927	16	1.4	286	125	BB115535	BB115535
c 928	16	1.4	286	128	BB237597	BB237597
c 929	16	1.4	286	128	BB255072	BB255072
c 930	16	1.4	286	130	BB305124	BB305124
c 931	16	1.4	286	135	BB521513	

High quality sequence start: 7

High quality sequence stop: 521.

Location/Qualifiers

## FEATURES

source

1. .522  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CR0267"  
/dev\_stage="Adult"  
/note="Organ: colon; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 123 a 133 c 123 g 143 t

## ORIGIN

Query Match 10.2%; Score 114; DB 120; Length 522;  
Best Local Similarity 100.0%; Pred. No. 1.1e-50;  
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 351 CATCAACTTAACATCGGGGATGCTGCTCCGACTTGCATTAGCATGGACGGTACATCGC 410  
|||||  
Db 57 CATCAACTTAACATCGGGGATGCTGCTCCGACTTGCATTAGCATGGACGGTACATCGC 116  
QY 411 CATTGTACAGCGCACTAAGTCATTCCGGCTCCGATCCAGAACACTACCGCGCAG 464  
|||||  
Db 117 CATTGTACAGCGCACTAAGTCATTCCGGCTCCGATCCAGAACACTACCGCGCAG 170

## RESULT 2

AI808029/c

LOCUS

DEFINITION wf52q04.x1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
IMAGE:2359254 3' similar to SW:CKR6\_HUMAN P51684 C-C CHEMOKINE  
RECEPTOR TYPE 6 ;, mRNA sequence.

ACCESSION AI808029

VERSION AI808029.1 GI:5394517

KEYWORDS EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 535 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

## FEATURES

source

1. .382  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:2359254"  
/clone\_lib="Soares\_NFL\_T\_GBC\_S1"  
/lab\_host="DH10B"  
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI-CGAP GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo. \*  
BASE COUNT 100 a 89 c 76 g 117 t

## ORIGIN

Query Match 7.1%; Score 79; DB 102; Length 382;  
Best Local Similarity 100.0%; Pred. No. 1.3e-31;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1041 TTCTCGGAGACCACTGAGCGCAGATACGACAAATGCTGCTCTTCACTATGATA 1100

|||||

Db 382 TTCTCGGAGACCACTGAGCGCAGATACGACAAATGCTGCTCTTCACTATGATA 323

|||||

QY 1101 GAAAGCTGAGTCTCCCTAA 1119

|||||

Db 322 GAAAGCTGAGTCTCCCTAA 304

## RESULT 3

AI045155/c

LOCUS

DEFINITION UI-R-C1-kk-c-08-0-UI.s1 UI-R-C1 Rattus norvegicus cDNA clone  
UI-R-C1-kk-c-08-0-UI 3', mRNA sequence.

ACCESSION AI045155

VERSION AI045155.1 GI:3291974

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized adult Spleen library. cDNA Library Preparation: M. Fatima Bonaldo, Ph.D. Clone distribution: clones will be available through Research Genetics This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID-1775435 The following repetitive elements were found in this cDNA sequence:

1-45, >POLY\_A#Simple\_repeat

Seq primer: M13 Forward

POLYA-No.

Location/Qualifiers

1. .461

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-C1-kk-c-08-0-UI"

/clone\_lib="UI-R-C1"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-C1 library is a subtracted library derived from the UI-R-CO library, which is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung,

brain, liver, kidney, heart, spleen, ovary, and muscle.

The UI-R-El library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-cl) was constructed as follows: PCR amplified cDNA inserts from UI-R-CO clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-CO library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-cl library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996).

BASE COUNT 100 a 113 c 108 g 138 t 2 others  
ORIGIN

Query Match 2.1%; Score 23; DB 15; Length 461;  
Best Local Similarity 100.0%; Pred. No. 0.44;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 880 GTCTGCTTCCTCGCACTGCTG 902  
|||||  
Db 325 GTCTGCTTCCTCGCACTGCTG 303

## RESULT 4

BF105117 822 bp mRNA EST 19-OCT-2000  
LOCUS 601822118F1 NIH\_MGC\_75 Homo sapiens cDNA clone IMAGE:4042301 5',  
DEFINITION mRNA sequence.

ACCESSION BF105117  
VERSION BF105117.1 GI:10887643  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE NTH-MGC http://mgc.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLC863 row: c column: 06

High quality sequence stop: 517.

## FEATURES

source

1. .822  
Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4042301"

/clone\_lib="NIH\_MGC\_75"

/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site:1;

Sfil (ggcgctcgcc); Site:2: Sfil (ggcattatggcc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGCCCAATTATGGCC-3' and 3' adaptor sequence:

5'-ATTCTAGAGCGCCGCGCCGACATG-dT(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.65

kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

BASE COUNT 205 a 213 c 196 g 208 t  
ORIGIN

Query Match 1.9%; Score 21; DB 144; Length 822;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 ACCAGTGAGACCGCAGATAC 1071

|||||

Db 698 ACCAGTGAGACCGCAGATAC 718

## RESULT 5

BE950816

LOCUS BE950816 396 bp mRNA EST 04-OCT-2000

DEFINITION UI-M-CE0-azd-d-01-0-UI-s1 NIH\_BMAP\_Ret3 Mus musculus cDNA clone

UI-M-CE0-azd-d-01-0-UI-3', mRNA sequence.

ACCESSION BE950816

VERSION BE950816.1 GI:10589524

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 396)

AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

## COMMENT

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the

retina tissue cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH

GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP

cDNAs whose availability will be considered under appropriate and

limited collaborative arrangements The following repetitive

elements were found in this cDNA sequence: 4-178, >B2#SINE/B2

181-306, >B1\_MM#SINE/Alu

Seq primer: M13 Forward

POLYA=yes

location/Qualifiers

1. .396

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-CE0-azd-d-01-0-UI"

/clone\_lib="NIH\_BMAP\_Ret3"

/dev\_stage="6 weeks"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site:1: Not I; Site:2: Eco RI; The

NIH\_BMAP\_Ret3 library is derived from mouse retina tissue.

For a detailed description of the library from which this

clone was derived, please visit our web site at

brainest.eng.uiowa.edu.

TAG\_LIB=NIH\_BMAP\_Ret3

TAG\_TISSUE=adult-retina

```

BASE COUNT      100 a      71 c      95 g      130 t
ORIGIN

Query Match      1.8%; Score 20; DB 141; Length 396;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 879 AGTCCTGGCTTCCTGCAC 898
|||||
Db 213 AGTCCTGGCTTCCTGCAC 232

RESULT 6
N83058/c
LOCUS
DEFINITION      TgSTz59g10.r1 TgRH Tachyzoite cDNA Toxoplasma gondii cDNA clone
ACCESSION      N83058
VERSION      TgSTz59g10.r1 5', mRNA sequence.
KEYWORDS      EST.
SOURCE      N83058.1 GI:1258811
ORGANISM      Toxoplasma gondii.
REFERENCE      1 (bases 1 to 412)
AUTHORS      Hehl, A., Manger, I., Marra, M., Sibley, L.D., Ajioke, J.A., Aslett, M.A., Dietrich, N., Dubucq, T., Hillier, L., Kucaba, T., Wan, K.L., Waterston, R.H., and Boothroyd, J.
JOURNAL      WashU-Merck EST Project
COMMENT      Contact: Marra M
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxo@wustl.wustl.edu
David Sibley at toxoest@borcim.wustl.edu for further information
relating to organism, clone or library availability.
Seq primer: T3
High quality sequence stop: 292.
Location/Qualifiers
1..412
/organism="Toxoplasma gondii"
/strain="RH"
/db_xref="taxon:5811"
/clone_lib="TgRH Tachyzoite cDNA"
/lab_host="XLI-Blue MRF"
/notes="Vector: Lambda ZAP; Site 1: EcoRI; Site 2: XhoI; Toxoplasma RH strain tachyzoites were grown in human foreskin fibroblast cultures in vitro. The library was synthesized from poly A RNAs by oligo d(T) priming and directionally cloned into the EcoRI to XhoI sites of the Lambda ZapII vector using the ZAP-cDNA synthesis kit (Stratagene). WARNING: the library contains a small percentage of cDNAs derived from the human host cells."
BASE COUNT      105 a      93 c      126 g      74 t      14 others
ORIGIN

Query Match      1.8%; Score 20; DB 155; Length 412;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 950 ACTTCTGAAGATCTTGAAG 969
|||||
Db 80 ACTTCTGAAGATCTTGAAG 61

TAG_SEQ=GTCCAGCGGCAC"

```

```

RESULT 7
BG044568
LOCUS
DEFINITION      saa30a06.y1 Gm-cl059 glycine max cDNA clone GENOME SYSTEMS CLONE
ACCESSION      BG044568
VERSION      BG044568.1 GI:12491430
KEYWORDS      EST.
SOURCE      soybean.
ORGANISM      Glycine max
REFERENCE      1 (bases 1 to 489)
AUTHORS      Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
JOURNAL      Public Soybean EST Project
COMMENT      Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
High quality sequence stop: 422.
Location/Qualifiers
1..489
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl059-396"
/clone_lib="Gm-cl059"
/tissue_type="Whole seedling, 2 week old, etiolated"
/lab_host="DH10B"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 2 week old etiolated whole seedlings of P1468916. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI- XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed in the laboratory of Dr. Randy Shoemaker at Iowa state university."
BASE COUNT      157 a      104 c      97 g      131 t
ORIGIN

Query Match      1.8%; Score 20; DB 172; Length 489;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 940 TTCAGAACTACTTCTGAA 959
|||||
Db 271 TTCAGAACTACTTCTGAA 290

RESULT 8
AZ023717/c
LOCUS
DEFINITION      AZ023717 513 bp DNA GSS
RPCI-23-385011.TJ RPCI-23 Mus musculus genomic clone RPCI-23-385011, DNA sequence.

```

```

ACCESSION      AZ023717
VERSION        AZ023717.1  GI:7099101
KEYWORDS       GSS.
SOURCE         house mouse.
ORGANISM       Mus musculus

REFERENCE
AUTHORS        Zhao,S., Nierman,W., Feldblum,T., Malek,J., Shatsman,S., Akinret
               B., Levins,M., Megann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
               and Fraser,C.M.
TITLE          Mouse BAC End Sequences from Library RPCI-23
JOURNAL        Unpublished (1999)
COMMENT        Other GSSs: RPCI-23-385011.TV
               Contact: Shaying Zhao
               Department of Eukaryotic Genomics
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850, USA
               Tel: 301 838 0200
               Fax: 301 838 0208
               Email: szhao@tigr.org
               Clones are derived from the mouse BAC library RPCI-23. For BAC
               library availability, please contact Pieter de Jong
               (pieter@dejong.med.buffalo.edu). Clones may be purchased from
               BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
               or from Resea ch Genetics (info@resgen.com). BAC end page:
               http://www.tigr.org/tdb/bac/ends/mouse/bac\_end\_intro.html
               Plate: 385 row: 0 column: 11
               Seq primer: SP6
               Class: BAC ends.

FEATURES
source         Location/Qualifiers
               1..513
               /organism="Mus musculus"
               /strain="C57BL/6J"
               /db_xref="taxon:10090"
               /clone="RPCI-23-385011"
               /clone_lib="RPCI-23"
               /sex="Female"
               /lab_host="DH10B"
               /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:
               EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
               brain genomic DNA was isolated and partially digested
               with a combination of EcoRI and EcoRI Methylase. Size
               selected DNA was cloned into the pBACE3.6 vector at the
               EcoRI sites. The ligation products were transformed into
               DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT     179 a 74 c 139 g 121 t
ORIGIN

Query Match    1.8%; Score 20; DB 215; Length 513;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 251 TCCTCTTTGTTCTTACTCTC 270
      |||||
Db 234 TCCTCTTTGTTCTTACTCTC 215

RESULT 9
AW733824/c
LOCUS          AW733824 530 bp mRNA EST 24-APR-2000
DEFINITION     SK84102.y1 Gm-cl035 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
               Gm-cl035-76.5, similar to TR:080731 080731 PUTATIVE
               PECTINACETYLESTERASE PRECURSOR. ;, mRNA sequence.
ACCESSION      AW733824
VERSION        AW733824.1  GI:7639503
KEYWORDS       EST.
SOURCE         soybean.
ORGANISM       Glycine max
               Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
               Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
               Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE
AUTHORS        Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
               Shoenberger,R., Keim,P., Vodkin,L., Erpelding,J., Corvelli,V., Khanna
               A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
               Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
               Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
               R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
               R., Waterston,R. and Wilson,R.
TITLE          Public Soybean EST Project
JOURNAL        Unpublished (1999)
COMMENT        Contact: Shoemaker R/Public Soybean EST Project
               Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: est@wustl.wustl.edu
               This clone is available through: Genome Systems, Inc. 4633 World
               Parkway Circle St. Louis, Missouri 63134 For further information
               call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
               427-3324 or contact: clones@genomesystems.com or
               info@genomesystems.com web site: www.genomesystems.com
               High quality sequence stop: 445.

FEATURES
source         Location/Qualifiers
               1..530
               /organism="Glycine max"
               /db_xref="taxon:3847"
               /clone="GENOME SYSTEMS CLONE ID: Gm-cl035-76"
               /clone_lib="Gm-cl035"
               /tissue_type="Immature leaves of greenhouse grown plants"
               /lab_host="DH10B"
               /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This
               cDNA library was constructed from mRNA isolated from
               immature leaves (unfurled trifoliolate) of greenhouse grown
               plants that were 2 weeks old. The library was prepared
               using the Life Technologies pSuperScript cDNA library
               construction kit. Complementary DNA was synthesized from
               mRNA using a poly (dT) sequence with a NotI restrictions
               site. SalI linkers adapters were ligated to the
               blunt-ended cDNA fragments followed by NotI digestion.
               The cDNA fragments were directionally cloned into the
               NotI-SalI restriction site of the pSPORT1 vector. The
               ligated cDNA fragments were transformed into E.coli
               ElectroMax DH10B host cells. This library was constructed
               in the laboratory of Dr. Lila Vodkin by Anu Khanna at the
               University of Illinois at Urbana-Champaign. e-mail:
               l-vodkin@uiuc.edu"
BASE COUNT     185 a 69 c 109 g 165 t 2 others
ORIGIN

Query Match    1.8%; Score 20; DB 119; Length 530;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 634 TTTGGTTTCTTTATCCCTTT 653
      |||||
Db 442 TTTGGTTTCTTTATCCCTTT 423

RESULT 10
AW773091
LOCUS          AW773091 568 bp mRNA EST 05-MAY-2000
DEFINITION     925002C05.y1 C. reinhardtii CC-2290, normalized, Lambda Zap II
               Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      AW773091
VERSION        AW773091.1  GI:7709602
KEYWORDS       EST.
SOURCE         Chlamydomonas reinhardtii.
ORGANISM       Chlamydomonas reinhardtii
               Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
               Chlamydomonadaceae; Chlamydomonas.
REFERENCE
AUTHORS        Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,

```

McDermott, J.P., Silflow, C., Stern, D. and Surzycki, R.  
 Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
 Unicellular System for Analyzing Gene Function and Regulation in  
 Vascular Plants; project phase 2  
 Unpublished (2000)  
 Contact: Elizabeth H. Harris  
 DCMB Box 91000  
 Duke University  
 Durham, NC 27708-1000, USA  
 Tel: 919 613 8164  
 Fax: 919 613 8177  
 Email: chlamy@duke.edu.

## FEATURES

source

1..568  
 /organism="Chlamydomonas reinhardtii"  
 /strain="CC-2290 wild type mt- S1 D2"  
 /db\_xref="taxon:3055"  
 /clone\_lib="C. reinhardtii CC-2290, normalized, Lambda Zap  
 II"  
 /note="Vector: pBluescript II SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; This library was constructed by John Davies and  
 Jeffrey McDermott. RNA was isolated from strain CC-2290  
 (Minnesota isolate of C. reinhardtii) grown to mid-log  
 phase in TAP (acetate containing) medium in the light.  
 PolyA mRNA was purified, and cDNA was synthesized and  
 directionally cloned into lambda ZAP II (Stratagene) in  
 the EcoRI (5') and XhoI (3') sites. pBluescript II SK-  
 plasmids were excised from the lambda ZAP clones by  
 superinfection with EXAssist (Stratagene) phage. The  
 library was normalized using method 4 described in Bonaldo  
 et al (1996) Genome Research 6: 791-806."  
 117 a 160 c 188 g 103 t

## BASE COUNT

ORIGIN

Query Match 1.8%; Score 20; DB 119; Length 568;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 444 ATCCGACACTACCGCGCA 463  
 |||||  
 Db 106 ATCCGACACTACCGCGCA 125  
 RESULT 11  
 AZ695745/c  
 LOCUS  
 DEFINITION  
 AZ695745 609 bp DNA GSS 24-JAN-2001  
 RPCI-23-240D14.TJ RPCI-23 Mus musculus genomic clone RPCI-23-240D14  
 , DNA sequence.  
 ACCESSION  
 AZ695745  
 VERSION  
 AZ695745.1 GI:12410181  
 GSS.  
 SOURCE  
 house mouse.  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 609)  
 Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret  
 , B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.  
 and Fraser, C.M.  
 Mouse BAC End Sequences from Library RPCI-23  
 Unpublished (1999)  
 Other\_GSS: RPCI-23-240D14.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCI-23. For BAC  
 library availability, please contact Pieter de Jong  
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end  
 page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 240 row: D column: 14  
 Seq primer: SP6  
 Class: BAC ends.

## FEATURES

source

Location/Qualifiers  
 1..609  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-240D14"  
 /clone\_lib="RPCI-23"  
 /sex="Female"  
 /lab\_host="DH10B"  
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site\_1:  
 EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBACe3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."  
 219 a 98 c 150 g 142 t

## BASE COUNT

ORIGIN

Query Match 1.8%; Score 20; DB 226; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 251 TCCTCTTTGTTCTTACTCTC 270  
 |||||  
 Db 339 TCCTCTTTGTTCTTACTCTC 320

## RESULT 12

CNS02TWL

LOCUS

DEFINITION

CNS02TWL 731 bp DNA GSS 15-MAY-2000

Tetraodon nigroviridis genome survey sequence T7 end of clone

164P10 of library G from Tetraodon nigroviridis, genomic survey

sequence.

ACCESSION

AL213726

VERSION

GI:7872545

KEYWORDS

GSS: genome survey sequence.

SOURCE

Tetraodon nigroviridis.

ORGANISM

Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Eurypterygii; Ctenosquamata; Acanthomorpha; Euacanthomorpha;

Holacanthopterygii; Acanthopterygii; Percomorpha;

Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

1 (bases 1 to 731)

Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizes, C., Fisher, C.,

Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and

Weissenbach, J.

Characterization and repeat analysis of the compact genome of the

freshwater pufferfish Tetraodon nigroviridis

Unpublished

2 (bases 1 to 731)

Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizes, C., Fisher, C.,

Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A.,

Saurin, W. and Weissenbach, J.

Human gene number estimate provided by genome wide analysis using

Tetraodon nigroviridis DNA sequence

Unpublished

3 (bases 1 to 731)

Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizes, C., Fisher, C.,

Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A.,

Saurin, W. and Weissenbach, J.

Direct Submission

Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis

genome. For more information, please take a look at

<http://www.genoscope.cns.fr/Tetraodon>.

Location/Qualifiers

```

source
1. .731
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="164p10"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG164DH05LP1-end : T7"
BASE COUNT 192 a 186 c 157 g 189 t 7 others
ORIGIN

Query Match 1.8%; Score 20; DB 230; Length 731;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 748 GTGGTCTGTTGTTCTGGC 767
|||||
DB 134 GTGGTCTGTTGTTCTGGC 153

RESULT 13
BE698296 260 bp mRNA EST 11-SEP-2000
LOCUS RC2-UT0021-070800-014-a04 UT0021 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE698296
VERSION BE698296.1 GI:10085456
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 260)
AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC2-UT0021-070
800-014-a04&t3=2000-08-07&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 260.
FEATURES
Location/Qualifiers
source 1. .260
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0021"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site1: SmaI;
Site2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 88 a 41 c 68 g 62 t 1 others
ORIGIN

```

```

Query Match 1.7%; Score 19; DB 138; Length 260;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 GGAATATTCTGTGTGTGAT 179
|||||
DB 7 GGAATATTCTGTGTGTGAT 25

RESULT 14
AQ413053/c 262 bp DNA GSS 23-MAR-1999
LOCUS RPCI-11-17502.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-17502,
DEFINITION DNA sequence.
ACCESSION AQ413053
VERSION AQ413053.1 GI:4472541
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 262)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@igr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pleter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: SP6
Class: BAC ends.
FEATURES
Location/Qualifiers
source 1. .262
/organism="Homo sapiens"
/db_xref="GDB:7567153"
/db_xref="taxon:9606"
/clone="RPCI-11-17502"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site1: EcoRI; Site2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT 62 a 54 c 74 g 69 t 3 others
ORIGIN

Query Match 1.7%; Score 19; DB 206; Length 262;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 TGCCACGTGCAAGTTGCTA 336
|||||
DB 34 TGCCACGTGCAAGTTGCTA 16

RESULT 15
AA245646 276 bp mRNA EST 10-MAR-1997
LOCUS mx01a03.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:678892 5'
DEFINITION similar to TR:G563829 G563829 FATTY ACID TRANSPORT PROTEIN. ;, mRNA
sequence.
ACCESSION AA245646

```

```

VERSION AA245646.1 GI:1876435
SOURCE EST.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 276)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicinep
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:418596
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 93.
FEATURES
Location/Qualifiers
1..276
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:678892"
/tissue_type="Liver"
/lab_host="DH10B"
/note="Vector: pTT3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCACAACTGAAGTGGAGCGCGCCGACATCTTTTCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
BASE COUNT 71 a 65 c 77 g 63 t
ORIGIN
Query Match 1.7%; Score 19; DB 4; Length 276;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 876 AGAAGTCTGGCTTCCCTG 894
|||||
Db 248 AGAAGTCTGGCTTCCCTG 230
RESULT 16
LOCUS AQ639171 280 bp DNA GSS 08-JUL-1999
DEFINITION 927P1-1H5-TV 927P1 Trypanosoma brucei genomic clone 927P1-1H5, DNA
sequence.
ACCESSION AQ639171
VERSION AQ639171.1 GI:5115881
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei.
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
1 (bases 1 to 280)
REFERENCE 1
AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S., Donelson,J.,
Fraser,C. and Adams,M.
TITLE Determination of clone end sequences from Trypanosoma brucei TREU
927/4 Pl library
JOURNAL Unpublished (1999)
COMMENT Other GSSs: 927P1-1H5-TP
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: nelsayed@tigr.org
For clone/filter availability, please contact Sara Melville
(sm160@mcle.bio.cam.ac.uk). Pl end sequences search page:
http://www.tigr.org/tdb/mdb/tbdb/.
Seq primer: T7
Class: Pl ends.
FEATURES
Location/Qualifiers
1..280
/organism="Trypanosoma brucei"
/strain="TREU927/4"
/db_xref="taxon:5691"
/clone="927P1-1H5"
/clone_lib="927P1"
/note="Vector: PAD10SACBII; Site_1: Bam HI; Constructed by
Sara Melville, University of Cambridge, UK and Nancy
Shepherd, Dupont Merck, Wilmington, DE. Genomic DNA was
isolated from Trypanosoma brucei (stock TREU927/4) and
partially digested with Sau 3AI. DNA fragments were cloned
into the Bam HI site of PAD10SACBII vector (Genbank
accession U09128). The average insert size is 65 Kb.
Coverage: approx 4.4 X the haploid non-minichromosomal
genome."
BASE COUNT 72 a 53 c 69 g 86 t
ORIGIN
Query Match 1.7%; Score 19; DB 209; Length 280;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 744 AGCTGTGGTCTGTGTGTTT 762
|||||
Db 16 AGCTGTGGTCTGTGTGTTT 34
RESULT 17
LOCUS AI374108 330 bp mRNA EST 19-JAN-1999
DEFINITION T6314 WYAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma
brucei rhodesiense cDNA 5' similar to invariant surface
glycoprotein 65 - Trypanosoma brucei, mRNA sequence.
ACCESSION AI374108
VERSION AI374108.1 GI:4164620
KEYWORDS EST.
SOURCE Trypanosoma brucei rhodesiense.
ORGANISM Trypanosoma brucei rhodesiense.
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
1 (bases 1 to 330)
REFERENCE 1
AUTHORS Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.
TITLE Generation of expressed sequence tags as physical landmarks in the
genome of Trypanosoma brucei
JOURNAL Unpublished (1996)
COMMENT Contact: Majiwa PAO
Molecular Biology Unit
International Livestock Research Institute
P.O. Box 30709, Nairobi, Kenya
Tel: 254-2 630743
Fax: 254-2 631499
Email: p.majiwa@cnet.com
Seq primer: T3 primer.
FEATURES
Location/Qualifiers
1..330

```

```

/organism="Trypanosoma brucei rhodesiense"
/db_xref="taxon:31286"
/clone_lib="MVAT4 bloodstream form of serodeme WRATatl.1"
/note="Vector: Lambda ZAP II (Stratagene); Site_1: EcoRI;
Site_2: XhoI; The mRNA was purified from a cloned
population of bloodstream trypanosomes reexpressing the
MVAT4 metacyclic variant surface glycoprotein (VSG). A
unidirectional oligo dt-primered EcoRI/XhoI cDNA library was
constructed in lambda ZAP II (Stratagene)."
BASE COUNT      93 a   54 c   83 g   100 t
ORIGIN
Query Match      1.7%; Score 19; DB 19; Length 330;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 744 AGCTGTGGTCTTGTTT 762
|||||
Db 118 AGCTGTGGTCTTGTTT 136

RESULT 18
AI374416      348 bp   mRNA      EST      19-JAN-1999
LOCUS      T6773 MVAT4 bloodstream form of serodeme WRATatl.1 Trypanosoma
DEFINITION brucei rhodesiense cDNA 5', mRNA sequence.
ACCESSION  AI374416
VERSION     AI374416.1 GI:4164928
KEYWORDS   EST.
SOURCE     Trypanosoma brucei rhodesiense.
ORGANISM   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1 (bases 1 to 348)
AUTHORS    Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.
TITLE      Generation of expressed sequence tags as physical landmarks in the
            genome of Trypanosoma brucei
COMMENT     Unpublished (1996)
            Contact: Majiwa PAO
            Molecular Biology Unit
            International Livestock Research Institute
            P.O. Box 30709, Nairobi, Kenya
            Tel: 254-2 630743
            Fax: 254-2 631499
            Email: p.majiwaecnet.com
            Seq primer: T3 primer.
            Location/Qualifiers
                1..348
                /organism="Trypanosoma brucei rhodesiense"
                /db_xref="taxon:31286"
                /clone_lib="MVAT4 bloodstream form of serodeme WRATatl.1"
                /note="Vector: Lambda ZAP II (Stratagene); Site_1: EcoRI;
                Site_2: XhoI; The mRNA was purified from a cloned
                population of bloodstream trypanosomes reexpressing the
                MVAT4 metacyclic variant surface glycoprotein (VSG). A
                unidirectional oligo dt-primered EcoRI/XhoI cDNA library was
                constructed in lambda ZAP II (Stratagene)."
BASE COUNT      100 a   59 c   87 g   102 t
ORIGIN
Query Match      1.7%; Score 19; DB 19; Length 348;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 744 AGCTGTGGTCTTGTTT 762
|||||
Db 96 AGCTGTGGTCTTGTTT 114

RESULT 19
AI707310

```

```

LOCUS      AI707310      355 bp   mRNA      EST      04-JUN-1999
DEFINITION T7429 MVAT4 bloodstream form of serodeme WRATatl.1 Trypanosoma
            brucei rhodesiense cDNA 5', mRNA sequence.
ACCESSION  AI707310
VERSION     AI707310.1 GI:4997086
KEYWORDS   EST.
SOURCE     Trypanosoma brucei rhodesiense.
ORGANISM   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1 (bases 1 to 355)
AUTHORS    Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.
TITLE      Generation of expressed sequence tags as physical landmarks in the
            genome of Trypanosoma brucei
COMMENT     Unpublished (1996)
            Contact: Majiwa PAO
            Molecular Biology Unit
            International Livestock Research Institute
            P.O. Box 30709, Nairobi, Kenya
            Tel: 254-2 630743
            Fax: 254-2 631499
            Email: p.majiwaecnet.com
            Seq primer: T3 primer.
            Location/Qualifiers
                1..355
                /organism="Trypanosoma brucei rhodesiense"
                /db_xref="taxon:31286"
                /clone_lib="MVAT4 bloodstream form of serodeme WRATatl.1"
                /note="Vector: Lambda ZAP II (Stratagene); Site_1: EcoRI;
                Site_2: XhoI; The mRNA was purified from a cloned
                population of bloodstream trypanosomes reexpressing the
                MVAT4 metacyclic variant surface glycoprotein (VSG). A
                unidirectional oligo dt-primered EcoRI/XhoI cDNA library was
                constructed in lambda ZAP II (Stratagene)."
BASE COUNT      94 a   62 c   89 g   110 t
ORIGIN
Query Match      1.7%; Score 19; DB 24; Length 355;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 744 AGCTGTGGTCTTGTTT 762
|||||
Db 115 AGCTGTGGTCTTGTTT 133

RESULT 20
AI464788      360 bp   mRNA      EST      15-MAR-2000
LOCUS      m280h09.y1 Soares mouse NML Mus musculus cDNA clone IMAGE:719777 5'
DEFINITION similar to TR:P97524 P97524 VERY-LONG-CHAIN ACYL-COA SYNTHETASE. ;,
            mRNA sequence.
ACCESSION  AI464788
VERSION     AI464788.1 GI:4318818
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 360)
AUTHORS    Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Willson,R.
            The WashU-NCI Mouse EST Project 1999
            Unpublished (1999)
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810

```

Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
This read is a RESEQUENCE of a previously sequenced mouse clone  
This read has been verified (found to hit its original self in the correct orientation)

MGI:445273

Seq primer: -40RP from Gibco

High quality sequence stop: 358

POLYA-No.

#### FEATURES

Location/Qualifiers

1..360

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone\_image="IMAGE:719777"

/clone\_lib="Soares mouse NML"

/tissue\_type="Liver"

/lab\_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAAGTGGAGCGCGCGAATCTTTTGTGTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

78 a 81 c 109 g 92 t

#### BASE COUNT

ORIGIN

Query Match 1.7%; Score 19; DB 20; Length 360;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 876 AGAAGTCCTGGCTTCCTG 894

|||||

Db 41 AGAAGTCCTGGCTTCCTG 23

#### RESULT 21

LOCUS

W84101 360 bp mRNA EST 12-AUG-1996  
T2925 MWAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma  
brucei rhodesiense cDNA 5', mRNA sequence.

ACCESSION

W84101

VERSION W84101.1 GI:1395242

KEYWORDS

EST

SOURCE

Trypanosoma brucei rhodesiense.

ORGANISM

Trypanosoma brucei rhodesiense

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

REFERENCE

1 (bases 1 to 360)

AUTHORS

Djikeng, A., Donelson, J.E. and Majiwa, P.A.O.

TITLE

Generation of expressed sequence tags as physical landmarks in the

genome of Trypanosoma brucei

Unpublished (1996)

JOURNAL

COMMENT

Contact: Majiwa PAO

Molecular Biology Unit

International Livestock Research Institute

P.O. Box 30709, Nairobi, Kenya

Tel: 254-2 630743

Fax: 254-2 631499

Email: p.majiwa@cgnnet.com

Seq primer: T3 primer.

Location/Qualifiers

1..360

/organism="Trypanosoma brucei rhodesiense"

/db\_xref="taxon:31286"

/clone\_lib="MWAT4 bloodstream form of serodeme WRATat1.1"

/note="Vector: Lambda ZAP II (Stratagene); Site\_1: EcoRI;

Site\_2: XhoI; The mRNA was purified from a cloned

population of bloodstream trypanosomes reexpressing the

FEATURES

source

MWAT4 metacyclic variant surface glycoprotein (VSG). A  
unidirectional oligo dt-primed EcoRI/XhoI cDNA library was  
constructed in lambda ZAP II (Stratagene)."  
BASE COUNT 95 a 63 c 82 g 120 t  
ORIGIN

Query Match 1.7%; Score 19; DB 158; Length 360;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 AGCTGTGCTGCTTGTGTTT 762

|||||

Db 62 AGCTGTGCTGCTTGTGTTT 80

#### RESULT 22

LOCUS

AI374268

DEFINITION

AI374268 362 bp mRNA EST 19-JAN-1999

T6480 MWAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma

brucei rhodesiense cDNA 5', mRNA sequence.

ACCESSION

AI374268

VERSION AI374268.1 GI:4164780

KEYWORDS

EST

SOURCE

Trypanosoma brucei rhodesiense.

ORGANISM

Trypanosoma brucei rhodesiense

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

REFERENCE

1 (bases 1 to 362)

AUTHORS

Djikeng, A., Donelson, J.E. and Majiwa, P.A.O.

TITLE

Generation of expressed sequence tags as physical landmarks in the

genome of Trypanosoma brucei

Unpublished (1996)

JOURNAL

COMMENT

Contact: Majiwa PAO

Molecular Biology Unit

International Livestock Research Institute

P.O. Box 30709, Nairobi, Kenya

Tel: 254-2 630743

Fax: 254-2 631499

Email: p.majiwa@cgnnet.com

Seq primer: T3 primer.

Location/Qualifiers

1..362

/organism="Trypanosoma brucei rhodesiense"

/db\_xref="taxon:31286"

/clone\_lib="MWAT4 bloodstream form of serodeme WRATat1.1"

/note="Vector: Lambda ZAP II (Stratagene); Site\_1: EcoRI;

Site\_2: XhoI; The mRNA was purified from a cloned

population of bloodstream trypanosomes reexpressing the

MWAT4 metacyclic variant surface glycoprotein (VSG). A

unidirectional oligo dt-primed EcoRI/XhoI cDNA library was

constructed in lambda ZAP II (Stratagene)."

BASE COUNT 99 a 59 c 92 g 112 t

ORIGIN

Query Match

Best Local Similarity 1.7%; Score 19; DB 19; Length 362;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 AGCTGTGCTGCTTGTGTTT 762

|||||

Db 116 AGCTGTGCTGCTTGTGTTT 134

#### RESULT 23

LOCUS

AI215388

DEFINITION

AI215388 367 bp mRNA EST 21-OCT-1998

T6273 MWAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma

brucei rhodesiense cDNA 5', mRNA sequence.

ACCESSION

AI215388

VERSION AI215388.1 GI:3778989

KEYWORDS

EST

SOURCE  
ORGANISM Trypanosoma brucei rhodesiense.  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma

REFERENCE  
1 (bases 1 to 367)  
AUTHORS Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.  
TITLE Generation of expressed sequence tags as physical landmarks in the genome of Trypanosoma brucei  
JOURNAL Unpublished (1996)  
COMMENT Contact: Majiwa PAO  
Molecular Biology Unit  
International Livestock Research Institute  
P.O. Box 30709, Nairobi, Kenya  
Tel: 254-2 630743  
Fax: 254-2 631499  
Email: p.majiwa@cgnnet.com  
Seq primer: T3 primer.  
Location/Qualifiers  
1..367  
/organism="Trypanosoma brucei rhodesiense"  
/db\_xref="taxon:31286"  
/clone\_lib="MVAT4 bloodstream form of serodeme WRATat1.1"  
/note="Vector: Lambda ZAP II (Stratagene); Site\_1: EcoRI; Site\_2: XhoI; The mRNA was purified from a cloned population of bloodstream trypanosomes reexpressing the MVAT4 metacyclic variant surface glycoprotein (VSG). A unidirectional oligo dT-primed EcoRI/XhoI cDNA library was constructed in lambda ZAP II (Stratagene)."

BASE COUNT 108 a 64 c 89 g 106 t  
ORIGIN

Query Match 1.7%; Score 19; DB 17; Length 367;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 744 AGCTGTGGTCTGTGTTT 762  
|||||  
Db 116 AGCTGTGGTCTGTGTTT 134

RESULT 24  
AI707317 370 bp mRNA EST 04-JUN-1999  
LOCUS T7437 MVAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma  
DEFINITION brucei rhodesiense cDNA 5', mRNA sequence.  
ACCESSION AI707317 GI:4997093  
VERSION AI707317  
KEYWORDS EST.  
SOURCE Trypanosoma brucei rhodesiense.  
ORGANISM Trypanosoma brucei rhodesiense.  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma.

REFERENCE  
1 (bases 1 to 370)  
AUTHORS Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.  
TITLE Generation of expressed sequence tags as physical landmarks in the genome of Trypanosoma brucei  
JOURNAL Unpublished (1996)  
COMMENT Contact: Majiwa PAO  
Molecular Biology Unit  
International Livestock Research Institute  
P.O. Box 30709, Nairobi, Kenya  
Tel: 254-2 630743  
Fax: 254-2 631499  
Email: p.majiwa@cgnnet.com  
Seq primer: T3 primer.  
Location/Qualifiers  
1..370  
/organism="Trypanosoma brucei rhodesiense"  
/db\_xref="taxon:31286"  
/clone\_lib="MVAT4 bloodstream form of serodeme WRATat1.1"  
/note="Vector: Lambda ZAP II (Stratagene); Site\_1: EcoRI; Site\_2: XhoI; The mRNA was purified from a cloned

population of bloodstream trypanosomes reexpressing the MVAT4 metacyclic variant surface glycoprotein (VSG). A unidirectional oligo dT-primed EcoRI/XhoI cDNA library was constructed in lambda ZAP II (Stratagene)."

BASE COUNT 100 a 58 c 94 g 118 t  
ORIGIN

Query Match 1.7%; Score 19; DB 24; Length 370;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 744 AGCTGTGGTCTGTGTTT 762  
|||||  
Db 119 AGCTGTGGTCTGTGTTT 137

RESULT 25  
AI510879 402 bp mRNA EST 15-MAR-1999  
LOCUS T6885 MVAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma  
DEFINITION brucei rhodesiense cDNA 5', mRNA sequence.  
ACCESSION AI510879  
VERSION AI510879.1 GI:4416580  
KEYWORDS EST.  
SOURCE Trypanosoma brucei rhodesiense.  
ORGANISM Trypanosoma brucei rhodesiense.  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma.

REFERENCE  
1 (bases 1 to 402)  
AUTHORS Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.  
TITLE Generation of expressed sequence tags as physical landmarks in the genome of Trypanosoma brucei  
JOURNAL Unpublished (1996)  
COMMENT Contact: Majiwa PAO  
Molecular Biology Unit  
International Livestock Research Institute  
P.O. Box 30709, Nairobi, Kenya  
Tel: 254-2 630743  
Fax: 254-2 631499  
Email: p.majiwa@cgnnet.com  
Seq primer: T3 primer.  
Location/Qualifiers  
1..402  
/organism="Trypanosoma brucei rhodesiense"  
/db\_xref="taxon:31286"  
/clone\_lib="MVAT4 bloodstream form of serodeme WRATat1.1"  
/note="Vector: Lambda ZAP II (Stratagene); Site\_1: EcoRI; Site\_2: XhoI; The mRNA was purified from a cloned population of bloodstream trypanosomes reexpressing the MVAT4 metacyclic variant surface glycoprotein (VSG). A unidirectional oligo dT-primed EcoRI/XhoI cDNA library was constructed in lambda ZAP II (Stratagene)."

BASE COUNT 111 a 70 c 92 g 129 t  
ORIGIN

Query Match 1.7%; Score 19; DB 21; Length 402;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 744 AGCTGTGGTCTGTGTTT 762  
|||||  
Db 116 AGCTGTGGTCTGTGTTT 134

RESULT 26  
AA014588/c 435 bp mRNA EST 31-JUL-1996  
LOCUS mg93a01.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA  
DEFINITION clone IMAGE:440520 5', similar to PIR:A55093 A55093 fatty acid transport protein precursor - mouse ;, mRNA sequence.  
ACCESSION AA014588

```

VERSION      AA014588.1  GI:1475468
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus

REFERENCE
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 435)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE        The WashU-HMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
              WashU-HMI Mouse EST Project
              Washington University School of MedicineP
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: mouseest@watson.wustl.edu
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:265856
              Seq primer: mob.REGA+ET
              High quality sequence stop: 405.
              Location/Qualifiers
                1. .435
                  /organism="Mus musculus"
                  /strain="C57BL/6J"
                  /db_xref="taxon:10090"
                  /clone="IMAGE:440520"
                  /clone_lib="Soares mouse embryo NBMEl3.5 14.5"
                  /sex="unknown"
                  /tissue_type="embryo"
                  /dev_stage="13.5-14.5dpc total fetus"
                  /lab_host="DH10B"
                  /note="vector: p7T3p-Pac (Pharmacia) with a modified
                  polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                  was primed with a Not I - oligo(dT) primer 15',
                  TGTTACCAATCTGAAGTGGGCGCGCGGAAATTTTTTTTTTTTTTTT
                  T 3', on equal amounts of mRNA from 2 13.5dpc and 2
                  14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
                  State Univ., from 2 ]; double-stranded cDNA was ligated to
                  Eco RI adaptors (Pharmacia), digested with Not I and
                  cloned into the Not I and Eco RI sites of the modified
                  pT7T3 vector. Library went through one round of
                  normalization, and was constructed by Bento Soares and
                  M.Fatina Bonaldo. "
                108 a 106 c 122 g 99 t
                BASE COUNT 108 a 106 c 122 g 99 t
                ORIGIN

Query Match 1.78; Score 19; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 876 AGAGTCTCTGCTTTCCTG 894
|||||
DB 329 AGAGTCTCTGCTTTCCTG 311

RESULT 27
LOCUS      TA390D01Q 455 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 390d01, reverse sequence,
            genomic survey sequence.
ACCESSION AL498801
VERSION   AL498801.1 GI:11874523
KEYWORDS  GSS.
SOURCE    Trypanosoma brucei.
ORGANISM  Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma
1 (bases 1 to 455)
Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
Melville,S.E., Rajandream,M.A. and Barrell,B.G.
Direct Submission
Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
nhl@sanger.ac.uk
Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (
4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.
Location/Qualifiers
  1. .455
    /organism="Trypanosoma brucei"
    /strain="TREU927"
    /db_xref="taxon:5691"
    /clone="390d01"
  139 a 80 c 113 g 123 t
  BASE COUNT 139 a 80 c 113 g 123 t
  ORIGIN

Query Match 1.7%; Score 19; DB 233; Length 455;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 AGCTGTGCTGTGTGTTT 762
|||||
DB 205 AGCTGTGCTGTGTGTTT 223

RESULT 28
LOCUS      AA254935 457 bp mRNA EST 14-MAR-1997
DEFINITION mz80h09.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:719777 5',
            similar to TR:G563829 G563829 FATTY ACID TRANSPORT PROTEIN. ;, mRNA
            sequence.
ACCESSION  AA254935
VERSION     AA254935.1 GI:1889521
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 457)
            Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
            The WashU-HMI Mouse EST Project
            Unpublished (1996)
            Contact: Marra M/Mouse EST Project
            WashU-HMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:445273
            Seq primer: -28ml3 rev2 ET from Amersham

```

High quality sequence stop: 453.

## FEATURES

source

1. .457  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:719933"  
/clone\_lib="Soares mouse NML"  
/tissue\_type="Liver"  
/lab\_host="DH10B"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5,  
TGTTACCAATCTGAAGTGGCGGCGGAATCTTTTTTTTTTTT 3'];  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5,  
TGTTACCAATCTGAAGTGGCGGCGGAATCTTTTTTTTTTTT 3'];  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

BASE COUNT 97 a 112 c 130 g 118 t

## ORIGIN

Query Match 1.7%; Score 19; DB 4; Length 457;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 876 AGAAGTCTGCTTCTCTG 894

Db 41 AGAAGTCTGCTTCTCTG 23

## RESULT 29

AA255136/c

LOCUS

DEFINITION AA255136 463 bp mRNA EST 14-MAR-1997  
m82f03.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:719933 5'  
similar to TR:G563829 G563829 FATTY ACID TRANSPORT PROTEIN. ;, mRNA  
sequence.

ACCESSION AA255136

VERSION AA255136.1 GI:1889755

KEYWORDS EST.

SOURCE house mouse.

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 463)  
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:445429

TITLE

JOURNAL

COMMENT

Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 399.  
Location/Qualifiers  
1. .463  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:719933"  
/clone\_lib="Soares mouse NML"  
/tissue\_type="Liver"  
/lab\_host="DH10B"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

## FEATURES

source

1. .463  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:719933"  
/clone\_lib="Soares mouse NML"  
/tissue\_type="Liver"  
/lab\_host="DH10B"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5,  
TGTTACCAATCTGAAGTGGCGGCGGAATCTTTTTTTTTTTT 3'];  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

BASE COUNT 109 a 111 c 131 g 112 t

## ORIGIN

Query Match 1.7%; Score 19; DB 4; Length 463;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 876 AGAAGTCTGCTTCTCTG 894

Db 443 AGAAGTCTGCTTCTCTG 425

## RESULT 30

AV397592

LOCUS

DEFINITION AV397592 468 bp mRNA EST 29-SEP-2000

CDNA clone CL02a01\_r, mRNA sequence.

ACCESSION AV397592

VERSION AV397592.1 GI:6551808

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

CONTACT: Yasukazu Nakamura

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

Location/Qualifiers

1. .468

/organism="Chlamydomonas reinhardtii"

/strain="C9"

/db\_xref="taxon:3055"

/clone="CL02a01\_r"

/clone\_lib="Chlamydomonas reinhardtii C9"

/dev\_stage="photoautotrophic growth"

/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:

XhoI"

BASE COUNT 114 a 156 c 85 g 113 t

ORIGIN

Query Match

Best Local Similarity

Matches

19; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy 890 TCCTGCACCTGCTGCTGAA 908

Db 38 TCCTGCACCTGCTGCTGAA 56

## RESULT 31

AI881128

LOCUS

DEFINITION

AI881128 481 bp mRNA EST 21-JUL-1999

T8077 MVAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma

brucei rhodesiense cDNA 5', mRNA sequence.

```

ACCESSION   A1881128
VERSION     A1881128.1  GI:55551177
KEYWORDS    EST.
SOURCE      Trypanosoma brucei rhodesiense.
ORGANISM    Trypanosoma brucei rhodesiense.
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 481)
AUTHORS    Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.
TITLE      Generation of expressed sequence tags as physical landmarks in the
            genome of Trypanosoma brucei
JOURNAL     Unpublished (1996)
COMMENT     Molecular Biology Unit
            International Livestock Research Institute
            P.O. Box 30709, Nairobi, Kenya
            Tel: 254-2 630743
            Fax: 254-2 631499
            Email: p.majiwa@cgiar.net
Seq primer: T3 primer.
            Location/Qualifiers
FEATURES             source
     1..481
     /organism="Trypanosoma brucei rhodesiense"
     /db_xref="taxon:31286"
     /clone_lib="MVA74 bloodstream form of serodene WRATat1.1"
     /note="Vector: Lambda ZAP II (Stratagene); Site_1: EcoRI;
     Site_2: XhoI; The mRNA was purified from a cloned
     population of bloodstream trypanosomes reexpressing the
     MVA74 metacyclic variant surface glycoprotein (VSG). A
     unidirectional oligo dt-primer EcoRI/XhoI cDNA library was
     constructed in lambda ZAP II (Stratagene)."
BASE COUNT   134 a   85 c   109 g   153 t
ORIGIN
Query Match      1.7%; Score 19; DB 103; Length 481;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  744  AGCTGCTGCTGCTGCTGTTT 762
        |||||||
Db  117  AGCTGCTGCTGCTGCTGTTT 135

RESULT  32
LOCUS   BF443898             484 bp      mRNA           EST           01-DEC-2000
DEFINITION  261560 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   BF443898
VERSION     BF443898.1  GI:11503990
KEYWORDS    EST.
SOURCE      pig.
ORGANISM    Sus scrofa
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 484)
AUTHORS    Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
TITLE      Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
JOURNAL     Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smithemail@marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR Primers
            FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTCCCGAGTCACGACG
Plate: 95 row: L column: 9
Seq primer: ATTTAGTGACACTATAG.
            Location/Qualifiers
FEATURES             source
     1..484
     /organism="Sus scrofa"
     /db_xref="taxon:9823"
     /clone_lib="MARC 2P1G"
     /tissue_type="pooled"
     /lab_host="DH10B"
     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
     Library made from pooled tissue from testis, ovary,
     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT   166 a   90 c   139 g   89 t
ORIGIN
Query Match      1.7%; Score 19; DB 148; Length 484;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  506  TCTCCAGCTCAACTTTTGT 524
        |||||||
Db  416  TCTCCAGCTCAACTTTTGT 398

RESULT  33
LOCUS   AZ646557             490 bp      DNA           GSS           14-DEC-2000
DEFINITION  1M0512F11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0512F11 R, DNA sequence.
ACCESSION   AZ646557
VERSION     AZ646557.1  GI:11777143
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 490)
AUTHORS    Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL     Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0512 row: F column: 11
            Seq primer: CACACAGGAACACGATGACC
            Class: plasmid ends
            High quality sequence stop: 490.
            Location/Qualifiers
FEATURES             source
     1..490
     /organism="Mus musculus"
     /strain="C57BL/6J"
     /db_xref="taxon:10090"
     /clone="UUGC1M0512F11"
     /clone_lib="Mouse 10kb plasmid UUGC1M library"
     /sex="Male"
     /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
     /note="Vector: PWD42nv; Purified genomic DNA from M.
            musculus C57BL/6J (male) was obtained from the Jackson
            Laboratory Mouse DNA Resource
            (http://www.jax.org/resources/documents/dnares/). The DNA
            was hydrodynamically sheared by repeated passage through a

```

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pW042 (gll4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 139 a 139 c 98 g 114 t  
ORIGIN

Query Match 1.7%; Score 19; DB 225; Length 490;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 CTTGTCTTACTCTCCCA 273  
|||||  
DB 414 CTTGTCTTACTCTCCCA 432

## RESULT 34

AI552452/c 494 bp mRNA EST 15-MAR-2000  
LOCUS m93a01.y1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA  
DEFINITION clone IMAGE:440520 5' similar to TR:P97524 P97524 VERY-LONG-CHAIN  
ACYL-COA SYNTHETASE. ; mRNA sequence.

ACCESSION AI552452  
VERSION AI552452.1 GI:4484815  
KEYWORDS EST.  
SOURCE house mouse.

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 494)  
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person  
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter  
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
Waterston, R. and Wilson, R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)  
CONTACT: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu

TITLE The WashU-NCI Mouse EST Project 1999  
JOURNAL Unpublished (1999)  
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
This read is a RESEQUENCE of a previously sequenced mouse clone  
This read has been verified (found to hit its original self in the  
correct orientation)  
MGI:265856  
Seq primer: -40RP from Gibco  
High quality sequence stop: 447  
POLYA-No.

## FEATURES

Location/Qualifiers  
1. .494  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:440520"  
/dev\_stage="13.5-14.5dpc mouse embryo NbME13.5 14.5"  
/sex="unknown"  
/tissue\_type="embryo"  
/dev\_stage="13.5-14.5dpc total fetus"

source

/lab\_host="DH108"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I oligo(dT) primer [5'  
TGTTCACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTTTT  
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2  
14.5dpc embryos (total RNA provided by Minoru KO, Wayne  
State Univ., from 2 l; double-stranded cDNA was ligated to  
Eco RI adaptors (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of the modified  
pT73 vector. Library went through one round of  
normalization, and was constructed by Bento Soares and  
M.Fatima Bonaldo."

BASE COUNT 122 a 120 c 138 g 114 t  
ORIGIN

Query Match 1.7%; Score 19; DB 21; Length 494;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 876 AGAAGTCCTGGCTTCTCTG 894  
|||||  
DB 331 AGAAGTCCTGGCTTCTCTG 313

## RESULT 35

AA060911/c 502 bp mRNA EST 23-SEP-1996  
LOCUS mj88b01.r1 Soares mouse p3NMf19.5 Mus musculus cDNA clone  
DEFINITION IMAGE:483145 5' similar to gb:X81579 M.musculus mRNA for  
insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION AA060911  
VERSION AA060911.1 GI:1554732  
KEYWORDS EST.  
SOURCE house mouse.

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 502)  
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
CONTACT: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu

TITLE The WashU-HHMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:293889  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 1.  
Location/Qualifiers  
1. .502  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:483145"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH108 (ampicillin resistant)"  
/note="Vector: pT73D (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTCACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTTTT  
double-stranded cDNA was size selected, ligated to Eco RI

## FEATURES

Location/Qualifiers  
1. .502  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:483145"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH108 (ampicillin resistant)"  
/note="Vector: pT73D (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTCACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTTTT  
double-stranded cDNA was size selected, ligated to Eco RI

source

adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

BASE COUNT 112 a 136 c 151 g 103 t  
 ORIGIN  
 Query Match 1.7%; Score 19; DB 1; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 591 GCCCATCAGGTGGAAGCTG 609  
 |||||  
 Db 264 GCCCATCAGGTGGAAGCTG 246

RESULT 36  
 A0003850/c  
 LOCUS  
 DEFINITION CIT-HSP-2288K15.TF CIT-HSP Homo sapiens genomic clone 2288K15, DNA  
 GSS 26-JUN-1998  
 ACCESSION A0003850  
 VERSION A0003850.1 GI:3081501  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 506)  
 Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuwa,H., Simon,M. and Venter,J.C.  
 TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)

JOURNAL  
 COMMENT Unpublished (1998)  
 Other\_GSSs: CIT-HSP-2288K15.TR  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)  
 Seq primer: M13-21;  
 Class: BAC ends.

FEATURES  
 source  
 Location/Qualifiers  
 1..506  
 /organism="Homo sapiens"  
 /db\_xref="GDB:7149978"  
 /db\_xref="taxon:9606"  
 /clone="2288K15"  
 /clone\_lib="CIT-HSP"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2: HindIII"  
 HindIII

BASE COUNT 162 a 61 c 98 g 185 t  
 ORIGIN  
 Query Match 1.7%; Score 19; DB 201; Length 506;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TTACTCAGTTGATTCTGAG 68  
 |||||  
 Db 57 TTACTCAGTTGATTCTGAG 39

RESULT 37  
 AI880998

LOCUS  
 DEFINITION T7766 MVAT4 bloodstream form of serodeme WRATat1.1 Trypanosoma brucei rhodesiense cDNA 5', mRNA sequence.  
 EST 21-JUL-1999

ACCESSION AI880998  
 VERSION AI880998.1 GI:5555047  
 KEYWORDS EST.  
 SOURCE Trypanosoma brucei rhodesiense.  
 ORGANISM Trypanosoma brucei rhodesiense.

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

REFERENCE  
 AUTHORS 1 (bases 1 to 519)  
 TITLE Djikeng,A., Donelson,J.E. and Majiwa,P.A.O.  
 COMMENT Generation of expressed sequence tags as physical landmarks in the genome of Trypanosoma brucei  
 Unpublished (1996)

JOURNAL  
 COMMENT Molecular Biology Unit  
 International Livestock Research Institute  
 P.O. Box 30709, Nairobi, Kenya  
 Tel: 254-2 630743  
 Fax: 254-2 631499  
 Email: p.majiwa@cgnet.com  
 Seq primer: T3 primer.

FEATURES  
 source  
 Location/Qualifiers  
 1..519

/organism="Trypanosoma brucei rhodesiense"  
 /db\_xref="taxon:31286"  
 /clone\_lib="MVAT4 bloodstream form of serodeme WRATat1.1"  
 /note="vector: Lambda ZAP II (Stratagene); Site\_1: EcoRI; Site\_2: XhoI; The mRNA was purified from a cloned population of bloodstream trypanosomes reexpressing the MVAT4 metacyclic variant surface glycoprotein (VSG). A unidirectional oligo dt-primer EcoRI/XhoI cDNA library was constructed in lambda ZAP II (Stratagene)."  
 156 a 88 c 118 g 157 t

BASE COUNT  
 ORIGIN  
 Query Match 1.7%; Score 19; DB 103; Length 519;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 AGCTGTGTGCTGTGTGTTT 762  
 |||||  
 Db 118 AGCTGTGTGCTGTGTGTTT 136

RESULT 38  
 AZ016585

LOCUS  
 DEFINITION RPCI-23-368N20.TV RPCI-23 Mus musculus genomic clone RPCI-23-368N20  
 GSS 25-FEB-2000  
 ACCESSION AZ016585  
 VERSION AZ016585.1 GI:7091969  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 534)

REFERENCE  
 AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.  
 TITLE Mouse BAC End Sequences from Library RPCI-23  
 JOURNAL Unpublished (1999)  
 COMMENT Other\_GSSs: RPCI-23-368N20.TJ  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@igrr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics ([info@resgen.com](mailto:info@resgen.com)). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 368 row: N column: 20  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source

Location/Qualifiers  
1. .534  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-368N20"  
/clone\_lib="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 162 a 107 c 71 g 194 t  
ORIGIN

Query Match 1.7%; Score 19; DB 215; Length 534;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 TCTTTGTTCTTACTCTCCC 272  
DB 109 TCTTTGTTCTTACTCTCCC 127

RESULT 39

BE033064 562 bp mRNA EST 09-JUL-2000  
LOCUS 133359 MARC lPIG Sus scrofa cDNA 5', mRNA sequence.  
DEFINITION BE033064  
ACCESSION BE033064  
VERSION BE033064.1 GI:8328064  
KEYWORDS EST.  
SOURCE pig.  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 562)  
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.  
Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine  
Unpublished (2000)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt.trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTTCCAGTCACGACG  
Plate: 67 row: H column: 15

TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
Location/Qualifiers  
1. .562  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC lPIG"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT 143 a 139 c 138 g 148 t  
ORIGIN

Query Match 1.7%; Score 19; DB 161; Length 568;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Seq primer: ATTTAGGTGACACTATAG.  
Location/Qualifiers  
1. .562  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC lPIG"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT 127 a 156 c 154 g 125 t  
ORIGIN

Query Match 1.7%; Score 19; DB 161; Length 562;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 597 CAGGTGGAAGCTGCTGATG 615  
DB 490 CAGGTGGAAGCTGCTGATG 508

RESULT 40

BE033241 568 bp mRNA EST 09-JUL-2000  
LOCUS 133608 MARC lPIG Sus scrofa cDNA 5', mRNA sequence.  
DEFINITION BE033241  
ACCESSION BE033241  
VERSION BE033241.1 GI:8328250  
KEYWORDS EST.  
SOURCE pig.  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 568)  
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.  
Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine  
Unpublished (2000)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt.trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTTCCAGTCACGACG  
Plate: 68 row: A column: 12  
Seq primer: ATTTAGGTGACACTATAG.

TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
Location/Qualifiers  
1. .568  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC lPIG"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT 143 a 139 c 138 g 148 t  
ORIGIN

Query Match 1.7%; Score 19; DB 161; Length 568;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT	42
BE468435/c	
LOCUS	
DEFINITION	
ACCESSION	

BASE COUNT 134 a 180 c 179 g 133 t  
 ORIGIN /clone\_lib="MAGE resequences, MAGD"  
 /note="vector: pbluescriptskm"

Query Match 1.7%; Score 19; DB 122; Length 626;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 765 GGCCTGTCAGATTCCTCAT 783  
 |||||||  
 Db 599 GGCCTGTCAGATTCCTCAT 617

RESULT 44  
 BF660122/c  
 LOCUS maa27h06.y1 631 bp mRNA EST 20-DEC-2000  
 DEFINITION similar to NCICGAP.Lil10 Mus musculus cDNA clone IMAGE:3812578 5'  
 PRECURSOR ;, mRNA sequence.

ACCESSION BF660122  
 VERSION BF660122.1 GI:11925256  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 631)  
 AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 TITLE NCICGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 COMMENT Tumor Gene Index  
 Unpublished (1997)

Other\_ESTS: maa27h06.xl  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCICGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

MGI:1454690  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 438.

FEATURES  
 source  
 1..631  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3812578"  
 /clone\_lib="NCICGAP\_Lil10"  
 /sex="female"  
 /dev\_stage="10 weeks"  
 /lab\_host="DH10B (TI phage-resistant)"  
 /note="Organ: liver; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.6 kb. Library constructed by Life  
 Technologies."

BASE COUNT 145 a 143 c 189 g 154 t  
 ORIGIN

Query Match 1.7%; Score 19; DB 151; Length 631;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 876 AGAAGTCTGGCTTCTCTG 894  
 |||||||  
 Db 288 AGAAGTCTGGCTTCTCTG 270

RESULT 45  
 AZ003038  
 LOCUS

DEFINITION AZ003038 633 bp DNA GSS 24-FEB-2000  
 RPCI-23-344M20.TV RPCI-23 Mus musculus genomic clone RPCI-23-344M20  
 , DNA sequence.

ACCESSION AZ003038  
 VERSION AZ003038.1 GI:7078394  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 633)  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret,  
 B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Kroi, M., de Jong, P.  
 and Fraser, C.M.  
 Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)  
 Other\_GSSs: RPCI-23-344M20.TJ  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)  
 or from Resea ch Genetics (info@resgen.com). BAC end page:  
 http://www.tigr.org/cdb/bac\_ends/mouse/bac\_end\_intro.html  
 Plate: 344 row: M column: 20  
 Seq primer: T7  
 Class: BAC ends.

FEATURES  
 source

1..633  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-344M20"  
 /clone\_lib="RPCI-23"  
 /sex="Female"  
 /lab\_host="DH10B"  
 /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site\_1:  
 EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBACE3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."  
 BASE COUNT 84 a 182 c 65 g 302 t  
 ORIGIN

Query Match 1.7%; Score 19; DB 214; Length 633;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 344 TCTATGCCATCAACTTAA 362  
 |||||||  
 Db 148 TCTATGCCATCAACTTAA 166

RESULT 46  
 AQ389111  
 LOCUS

DEFINITION AQ389111 640 bp DNA GSS 21-MAY-1999  
 RPCI11-157N16.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-157N16,  
 DNA sequence.

ACCESSION AQ389111  
 VERSION AQ389111.1 GI:4360134  
 KEYWORDS GSS.

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 640)  
 AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter  
 ,J.C.  
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 JOURNAL Unpublished (1997)  
 COMMENT Other\_GSSs: RPCI11-157N16.TV  
 Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
 Research Genetics (info@resgen.com). BAC end search page:  
 http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
 Seq primer: SP6  
 Class: BAC ends.

FEATURES  
 source Location/Qualifiers  
 1..640  
 /organism="Homo sapiens"  
 /db\_xref="GDB:7560231"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-157N16"  
 /clone\_lib="RPCI-11"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /note="vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPCI11 Human Male BAC Library"  
 BASE COUNT 199 a 112 c 116 g 213 t  
 ORIGIN

Query Match 1.7%; Score 19; DB 205; Length 640;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GATTATTTTGTCTAGTCA 39  
 |||||  
 Db 244 GATTATTTTGTCTAGTCA 262  
 RESULT 47  
 AQ379360  
 LOCUS 724 bp DNA GSS 20-MAY-1999  
 DEFINITION RPCI11-16A016.TV RPCI-11 Homo sapiens genomic clone RPCI-11-16A016,  
 DNA sequence.  
 ACCESSION AQ379360  
 VERSION AQ379360.1 GI:4350383  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 724)  
 AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter  
 ,J.C.  
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 JOURNAL Unpublished (1997)  
 COMMENT Other\_GSSs: RPCI11-16A016.TJ  
 Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
 Research Genetics (info@resgen.com). BAC end search page:  
 http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
 Seq primer: T7  
 Class: BAC ends.

FEATURES  
 source Location/Qualifiers  
 1..724  
 /organism="Homo sapiens"  
 /db\_xref="GDB:7562943"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-16A016"  
 /clone\_lib="RPCI-11"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /note="vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPCI11 Human Male BAC Library"  
 BASE COUNT 224 a 131 c 137 g 232 t  
 ORIGIN

Query Match 1.7%; Score 19; DB 205; Length 724;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GATTATTTTGTCTAGTCA 39  
 |||||  
 Db 244 GATTATTTTGTCTAGTCA 262  
 RESULT 48  
 AI047830  
 LOCUS 785 bp mRNA EST 08-JUL-1998  
 DEFINITION ud64b12.x1 Sugano mouse liver mlia Mus musculus cDNA clone  
 IMAGE:1450655 3' similar to TR:P97524 P97524 VERY-LONG-CHAIN  
 ACYL-COA SYNTHETASE. ;, mRNA sequence.  
 ACCESSION AI047830  
 VERSION AI047830.1 GI:3296117  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.  
 REFERENCE 1 (bases 1 to 785)  
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
 Waterston,R.  
 TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT Contact: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:923971  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: custom primer used  
 High quality sequence stop: 479.

FEATURES  
 source Location/Qualifiers  
 1..785  
 /organism="Mus musculus"  
 /strain="C57BL"

```

/db_xref="taxon:10090"
/clone="IMAGE:1450655"
/clone_lib="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
(note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
(CACCATGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGCGCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [CTGTGCCCTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTAAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGCAC."
BASE COUNT      191 a      212 c      185 g      196 t      1 others
ORIGIN

Query Match      1.7%; Score 19; DB 15; Length 785;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 876 AGAGCTCTGGCTTCTCTG 894
|||||
Db 724 AGAGCTCTGGCTTCTCTG 742
|||||

RESULT 49
BF235138/c      788 bp      mRNA      EST      14-NOV-2000
LOCUS
DEFINITION      602026409F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4161674 5',
mRNA sequence.
ACCESSION      BF235138.1 GI:11147372
VERSION
KEYWORDS
SOURCE          house mouse.
ORGANISM        Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 788)
Marra.M., Hillier.L., Kucaba.T., Martin.J., Beck.C., Wylie.T.,
Underwood.K., Steptoe.M., Theising.B., Allen.M., Bowers.Y., Person
.B., Swaller.T., Gibbons.M., Pape.D., Harvey.N., Schurk.R., Ritter
.E., Kohn.S., Shin.T., Jackson.Y., Cardenas.M., McCann.R.,
Waterston.R. and Willson.R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone.
This read has been verified (found to hit its original self in the
correct orientation)
MGI:224659
Seq primer: -40RP from Gibco
High quality sequence stop: 465
POLYA=No.
FEATURES
source
Location/Qualifiers
1..802
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:352859"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/Note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGCGGAATTTTCTTTTCTTTTCTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 ]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT      173 a      191 c      233 g      190 t      1 others
ORIGIN

```

```

ORIGIN

Query Match      1.7%; Score 19; DB 145; Length 788;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 876 AGAGCTCTGGCTTCTCTG 894
|||||
Db 534 AGAGCTCTGGCTTCTCTG 516
|||||

RESULT 50
AI430701/c      802 bp      mRNA      EST      15-MAR-2000
LOCUS
DEFINITION      mc59h06.y1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
clone IMAGE:352859 5' similar to gb:X15313 Mouse MPO mRNA for
myeloperoxidase (MOUSE);, mRNA sequence.
ACCESSION      AI430701
VERSION
KEYWORDS
SOURCE          house mouse.
ORGANISM        Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 802)
Marra.M., Hillier.L., Kucaba.T., Martin.J., Beck.C., Wylie.T.,
Underwood.K., Steptoe.M., Theising.B., Allen.M., Bowers.Y., Person
.B., Swaller.T., Gibbons.M., Pape.D., Harvey.N., Schurk.R., Ritter
.E., Kohn.S., Shin.T., Jackson.Y., Cardenas.M., McCann.R.,
Waterston.R. and Willson.R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone.
This read has been verified (found to hit its original self in the
correct orientation)
MGI:224659
Seq primer: -40RP from Gibco
High quality sequence stop: 465
POLYA=No.
FEATURES
source
Location/Qualifiers
1..802
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:352859"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/Note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGCGGAATTTTCTTTTCTTTTCTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 ]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT      187 a      240 c      203 g      172 t
ORIGIN

```